

Dhinaharan Nagamalai
Eric Renault
Murugan Dhanuskodi (Eds.)

Communications in Computer and Information Science

205

Advances in Digital Image Processing and Information Technology

First International Conference on Digital Image Processing
and Pattern Recognition, DPPR 2011
Tirunelveli, Tamil Nadu, India, September 2011
Proceedings

Communications
in Computer and Information Science

205

Dhinaharan Nagamalai Eric Renault
Murugan Dhanuskodi (Eds.)

Advances in Digital Image Processing and Information Technology

First International Conference on Digital Image
Processing and Pattern Recognition, DPPR 2011
Tirunelveli, Tamil Nadu, India, September 23-25, 2011
Proceedings

Volume Editors

Dhinaharan Nagamalai
Wireilla Net Solutions PTY Ltd
Melbourne, VIC, Australia
E-mail: dhinthia@yahoo.com

Eric Renault
Institut Telecom/Telecom SudParis (ex. GET-INT)
Departement Reseaux et Services Multimedia Mobiles (RS2M)
Samovar UMR INT-CNRS 5157
9, rue Charles Fourier, 91011 Evry Cedex, France
E-mail: eric.renault@it-sudparis.eu

Murugan Dhanuskodi
Manonmaniam Sundaranar University
Department of Computer Science and Engineering
Tirunelveli, Tamil Nadu, India
E-mail: dhanushkodim@yahoo.com

ISSN 1865-0929
ISBN 978-3-642-24054-6
DOI 10.1007/978-3-642-24055-3
Springer Heidelberg Dordrecht London New York

e-ISSN 1865-0937
e-ISBN 978-3-642-24055-3

Library of Congress Control Number: Applied for

CR Subject Classification (1998): I.2, I.4, H.3, I.5, H.4, C.2, K.6.5

© Springer-Verlag Berlin Heidelberg 2011

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

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.

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

Printed on acid-free paper

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

Preface

The First International Conference on Computer Science, Engineering and Information Technology (CCSEIT-2011), The First International Conference on Parallel, Distributed Computing Technologies and Applications (PDCTA-2011), and The First International Conference on Digital Image Processing and Pattern Recognition (DPPR-2011) were held in Tirunelveli - Tamil Nadu, India, during September 23–25, 2011. The events attracted many local and international delegates, presenting a balanced mixture of intellects from all over the world. The goal of this conference series is to bring together researchers and practitioners from academia and industry to focus on understanding parallel, distributed computing technologies, digital image processing and pattern recognition and all areas of computer science, information technology, and to establish new collaborations in these areas.

The CCSEIT 2011, PDCTA 2011 and DPPR 2011 committees invited original submissions from researchers, scientists, engineers, and students that illustrate research results, projects, survey works, and industrial experiences describing significant advances in the areas related to the relevant themes and tracks of the conferences. This effort guaranteed submissions from an unparalleled number of internationally recognized top-level researchers. All the submissions underwent a strenuous peer-review process which comprised expert reviewers. Besides the members of the Technical Program Committee, external reviewers were invited on the basis of their specialization and expertise. The papers were reviewed based on their technical content, originality, and clarity. The entire process, which includes the submission, review, and acceptance processes, was done electronically. All these efforts undertaken by the Organizing and Technical Program Committees led to an exciting, rich, and high-quality technical conference program, which featured high-impact presentations for all attendees to enjoy and to expand their expertise in the latest developments in this field.

There were a total 1,256 submissions to the conference, and the Technical Program Committee selected 185 papers for presentation at the conference and subsequent publication in the proceedings. This small introduction would be incomplete without expressing our gratitude and thanks to the General and Program Chairs, members of the Technical Program Committees, and external reviewers for their excellent and diligent work. Thanks to Springer for the strong support. Finally, we thank all the authors who contributed to the success of the conference. We also sincerely wish that all attendees benefited academically from the conference and wish them every success in their research.

Dhinaharan Nagamalai
Eric Renault
Murugan Dhanushkodi

Organization

General Chairs

David C. Wyld	Southeastern Louisiana University, USA
Michal Wozniak	Wroclaw University of Technology, Poland

Steering Committee Chairs

Murugan Dhanuskodi	Manonmaniam Sundaranar University, India
Jan Zizka	SoNet/DI, FBE, Mendel University in Brno, Czech Republic
John Karamitsos	University of the Aegean, Samos, Greece
Khoa N. Le	University of Western Sydney, Australia
Nabendu Chaki	University of Calcutta, India
Salah S. Al-Majeed	University of Essex, UK
Dhinaharan Nagamalai	Wireilla Net Solutions, Australia

Publicity Chairs

Roberts Masillamani	Hindustan University, India
Chih-Lin Hu	National Central University, Taiwan

Program Committee

A. Arokiasamy	Eastern Mediterranean University, Cyprus
A.P. Sathish Kumar	PSG Institute of Advanced Studies, India
Abdul Aziz	University of Central Punjab, Pakistan
Abdul Kadir Ozcan	The American University, Cyprus
Al-Sakib Khan Pathan	Kyung Hee University, South Korea
Andreas Rienner	Johannes Kepler University Linz, Austria
Andy Seddon	Asia Pacific Institute of Information Technology, Malaysia
Antelin Vijjila	Manonmaniam Sundaranar University, India
Arvinth Kumar	M.S. University, India
Athanasios Vasilakos	University of Western Macedonia, Greece
Atilla Elci	Eastern Mediterranean University, Cyprus
B. Srinivasan	Monash University, Australia

VIII Organization

Balasubramaniam	Manonmaniam Sundaranar University, India
Bong-Han Kim	Chongju University , South Korea
Boo-Hyung Lee	KongJu National University, South Korea
Brajesh Kumar Kaushik	Indian Institute of Technology, India
Charalampos Z. Patrikakis	National Technical University of Athens, Greece
Chih-Lin Hu	National Central University, Taiwan
Chin-Chih Chang	Chung Hua University, Taiwan
Cho Han Jin	Far East University, South Korea
Cynthia Dhinakaran	Hannam University, South Korea
Danda B. Rawat	Old Dominion University, USA
David W. Deeds	Shingu College, South Korea
Debasis Giri	Haldia Institute of Technology, India
Deepak Garg	Thapar University, India
Dhinaharan Nagamalai	Wireilla Net Solutions Pty Ltd., Australia
Dimitris Kotzinos	Technical Educational Institution of Serres, Greece
Dong Seong Kim	Duke University, USA
Emmanuel Bouix	iKlax Media, France
Eric Renault	Institut Telecom – Telecom SudParis, France
Farhat Anwar	International Islamic University , Malaysia
Firkhan Ali Bin Hamid Ali	Universiti Tun Hussein Onn Malaysia, Malaysia
Ford Lumban Gaol	University of Indonesia
Geuk Lee	Hannam University, South Korea
Girija Chetty	University of Canberra, Australia
H.V. Ramakrishnan	Dr. MGR University, India
Henrique Joao Lopes Domingos	University of Lisbon, Portugal
Ho Dac Tu	Waseda University, Japan
Hoang, Huu Hanh	Hue University, Vietnam
Hwangjun Song	Pohang University of Science and Technology, South Korea
J.Arunadevi	Thiagarajar College, Madurai, India
Jacques Demerjian	Communication & Systems, Homeland Security, France
Jae Kwang Lee	Hannam University, South Korea
Jan Zizka	SoNet/DI, FBE, Mendel University in Brno, Czech Republic
Jansirani	Manonmaniam Sundaranar University, India
Jeong-Hyun Park	Electronics Telecommunication Research Institute, South Korea
Jivesh Govil	Cisco Systems Inc. - CA, USA
Johann Groschd	University of Bristol, UK

John Karamitsos	University of the Aegean, Samos, Greece
Johnson Kuruvila	Dalhousie University, Halifax, Canada
Jose Enrique Armendariz-Inigo	Universidad Publica de Navarra, Spain
Jung-Gil Song	Hannam University, South Korea
Jungwook Song	Konkuk University, South Korea
K.P. Thooyamani	Bharath University, India
Kamaljit I Lakhtaria	Atmiya Institute of Technology & Science, India
Kannan	Anna University, Chennai, India
Khamish Malhotra	University of Glamorgan, UK
Khoa N. Le	Griffith University, Australia
Krzysztof Walkowiak	Wroclaw University of Technology, Poland
L. Ganesan	Alagappa University, India
Lu S. Veiga	Technical University of Lisbon, Portugal
Lu Yan	University of Hertfordshire, UK
Maode Ma	Nanyang Technological University, Singapore
Marco Rocchetti	Universty of Bologna, Italy
Michael Peterson	University of Hawaii at Hilo, USA
Michal Wozniak	Wroclaw University of Technology, Poland
Mohsen Sharifi	Iran University of Science and Technology, Iran
Murugan D.	Manonmaniam Sundaranar University, India
Murugeswari	M.S. University, India
Muthulakshmi	M.S. University, India
N. Krishnan	Manonmaniam Sundaranar University, India
Nabendu Chaki	University of Calcutta, India
Natarajan Meghanathan	Jackson State University, USA
Neerajkumar	Madha Vaisnavi Devi University, India
Nicolas Sklavos	Technological Educational Institute of Patras, Greece
Nidaa Abdual Muhsin Abbas	University of Babylon, Iraq
Paul D. Manuel	Kuwait University, Kuwait
Phan Cong Vinh	London South Bank University, UK
Ponpit Wongthongtham	Curtin University of Technology, Australia
Prabu Dorairaj	NetApp Inc., India
Rajalakshmi	Manonmaniam Sundaranar University, India
Rajendra Akerkar	Technomathematics Research Foundation, India
Rajesh Kumar P.	The Best International, Australia
Rajesh	Manonmaniam Sundaranar University, India
Rajesh Bawa	Punjabi University, India

Rajkumar Kannan	Bishop Heber College, India
Rakhesh Singh Kshetrimayum	Indian Institute of Technology-Guwahati, India
Ramayah Thurasamy	Universiti Sains Malaysia, Malaysia
Rituparna Chaki	West Bengal University of Technology, India
Roberts Masillamani	Hindustan University, India
S. Arumugam	Nandha Engineering College, India
S. Hariharan	B.S.Abdur Rahman University, India
Sadasivam	Manonmaniam Sundaranar University, India
Sagarmay Deb	Central Queensland University, Australia
Sajid Hussain	Acadia University, Canada
Sajid Hussain	Fisk University, USA
Salah S. Al-Majeed	University of Essex, UK
Sanguthevar Rajasekaran	University of Connecticut, USA
Sarmistha Neogy	Jadavpur University, India
Sattar B. Sadkhan	University of Babylon, Iraq
Seemabawa	Thapar University, India
Sergio Ilarri	University of Zaragoza, Spain
Serguei A. Mokhov	Concordia University, Canada
Seungmin Rho	Carnegie Mellon University, USA
Shivan Haran	Arizona State University, USA
Somitra Sanadhya	IIT-Delhi, India
Soodamani, ramalingam	University of Hertfordshire, UK
Sriman Narayana Iyengar	VIT University, India
Subha	M.S. University, India
Sudip Misra	Indian Institute of Technology-Kharagpur, India
Sundarapandian Viadyanathan	Vel Tech Dr.RR & Dr.SR Technical University, India
Sundaesan	Manonmaniam Sundaranar University, India
SunYoung Han	Konkuk University, South Korea
Suruliandi	Manonmaniam Sundaranar University, India
Susana Sargento	University of Aveiro, Portugal
Syed Rahman	University of Hawaii-Hilo, USA
Syed Rizvi	University of Bridgeport, USA
Taruna S.	Banasthali University, India
Thamaraiselvi	Madras Institute of Technology, India
Thambidurai	Pondicherry University, India
Velmurugan Ayyadurai	Center for Communication Systems, UK
Vishal Sharma	Metanoia Inc., USA
Wei Jie	University of Manchester, UK

Yan Luo	University of Massachusetts Lowell, USA
Yannick Le Moullec	Aalborg University, Denmark
Yao-Nan Lien	National Chengchi University, Taiwan
Yeong Deok Kim	Woosong University, South Korea
Yuh-Shyan Chen	National Taipei University, Taiwan
Yung-Fa Huang	Chaoyang University of Technology, Taiwan

External Reviewers

Alejandro Regalado Mendez	Universidad del Mar, Mexico
Alireza Mahini	Islamic Azad University-Gorgan, Iran
Amandeep Singh Thethi	Guru Nanak Dev University Amritsar, India
Ashok Kumar Sharma	YMCA Institute of Engineering, India
Ayman Khalil	Institute of Electronics and Telecommunications of Rennes (IETR), France
Buket Barkana	University of Bridgeport, USA
Christos Politis	Kingston University, UK
Hao Shi	Victoria University, Australia
Indrajit Bhattacharya	Kalyani Government Engineering College, India
Jyotirmay Gadewadikar	Alcorn State University, USA
Khoa N. Le	University of Western Sydney, Australia
Laili Almazaydeh	University of Bridgeport, USA
Lakshmi Rajamani	Osmania University, India
Michel Owayjan	American University of Science & Technology - AUST, Lebanon
Mohamed Hassan	American University of Sharjah, UAE
Monika Verma	Punjab Technical University, India
N.K. Choudhari	Bhagwati Chaturvedi College of Engineering, India
Nitiket N. Mhala	B.D. College of Engineering - Sewagram, India
Nour Eldin Elmadany	Arab Academy for Science and Technology, Egypt
Premanand K. Kadbe	Vidya Pratishthan's College of Engineering, India
R. Murali	Dr. Ambedkar Institute of Technology, Bangalore, India
Raman Maini	Punjabi University, India
Rushed Kanawati	LIPN - Université Paris 13, France
S.A.V. Satyamurty	Indira Gandhi Centre for Atomic Research, India
Shrikant K. Bodhe	Bosh Technologies, India

Sridharan
Utpal Biswas
Wichian Sittiprapaporn

CEG Campus - Anna University, India
University of Kalyani, India
Mahasarakham University, Thailand

Technically Supported By

Networks & Communications Community (NCC)
Digital Signal & Image Processing Community (DSIPC)
Computer Science & Information Technology Community (CSITC)

Organized By



ACADEMY & INDUSTRY RESEARCH COLLABORATION CENTER (AIRCC)

Table of Contents

Digital Image Processing and Pattern Recognition

Fuzzy Soft Thresholding Based Hybrid Denoising Model	1
<i>Sudipta Roy, Nidul Sinha, and Asoke Kr. Sen</i>	
Content-Based Image Retrieval Using a Quick SVM-Binary Decision Tree – QSVMBDT	11
<i>I. Felci Rajam and S. Valli</i>	
Implementation of LabVIEW Based Intelligent System for Speed Violated Vehicle Detection	23
<i>G. Adinarayana, B. Lakshmi Sirisha, K. Sri Rama Krishna, and M. Kantikiran</i>	
Feature Extraction from Lunar Images	34
<i>V. Tamililakkiya and K. Vani</i>	
A Hybrid Approach to Estimate True Density Function for Gene Expression Data	44
<i>Ganesh Kumar Pugalendhi, Mahibha David, and Aruldoss Albert Victoire</i>	
Top Down Hierarchical Histogram Based Approach for Printed Devnagri Script Character Isolation	55
<i>Trusha Gajjar, Rekha Teraiya, Gunvantsinh Gohil, and Mahesh Goyani</i>	
Unsupervised Medical Image Classification Based on Skew Gaussian Mixture Model and Hierarchical Clustering Algorithm	65
<i>Nagesh Vadaparathi, Srinivas Yarramalle, and P. Suresh Varma</i>	
An Analog VLSI Implementation for Cepstral Technique for Disparity Estimation in Stereoscopic Vision	75
<i>Harshit Agarwal, Sheena Sharma, and Chota Markan</i>	
Finding Optimal Set of Orthogonal Polynomial Operators for Efficient Texture Feature Extraction	86
<i>R. Suguna and P. Anandhakumar</i>	
Design of Optimized Neuro-Wavelet Based hybrid Model for Image Compression	96
<i>Deepak Gambhir, Navin Rajpal, and Vipula Singh</i>	

Writer Bsed Handwritten Document Image Retrieval Using Contourlet Transform	108
<i>M.S. Shirdhonkar and Manesh B. Kokare</i>	
Semantic Learning in Interactive Image Retrieval	118
<i>Pushpa B. Patil and Manesh Kokare</i>	
A New Method for Lossless Image Compression Using Recursive Crack Coding	128
<i>T. Meyyappan, S.M. Thamarai, and N.M. Jeya Nachiaban</i>	
A Study on Block Matching Algorithms and Gradient Based Method for Motion Estimation in Video Compression	136
<i>Deepa Mary Thomas</i>	
Automatic, Robust Face Detection and Recognition System for Surveillance and Security Using LabVIEW (sCUBE)	146
<i>C. Lakshmi Deepika, M. Alagappan, A. Kandaswamy, H. Wassim Ferose, and R. Arun</i>	
Global Chaos Synchronization of Hyperchaotic Lorenz Systems by Sliding Mode Control	156
<i>Sundarapandian Vaidyanathan and Sivaperumal Sampath</i>	
Segmentation of Micro Calcification Clusters in Digital Mammograms Using UIQI	165
<i>Patnala S.R. Chandra Murty, T. Sudheer, and E. Sreenivasa Reddy</i>	
A Denoising Approach for Salt and Pepper Noise Corrupted Image at Higher Noise Density	173
<i>Debika Dey, Samanwita Laha, Shalini Chowdhury, and Subhojit Sarker</i>	
Blind Medical Image Watermarking Technique for Secure Recovery of Hidden Data	185
<i>Shaik Basheera, D. Bhanu Prakash, and P.V. Naganjaneyulu</i>	
Global Chaos Synchronization of Lü and Pan Systems by Adaptive Nonlinear Control	193
<i>Sundarapandian Vaidyanathan and Karthikeyan Rajagopal</i>	
Reducing False Positives in Minutia Detection by Using the Proposed Fingerprint Alignment Technique	203
<i>P. Jaganathan and M. Rajinikannan</i>	
Modeling and Simulation of Respiratory Controller Using Labview	212
<i>P. Srinivas, V. Naga Prudhvi Raj, and S. Rajesh</i>	

A Comparative Study on Human Spermatozoa Images Classification with Artificial Neural Network Based on FOS, GLCM and Morphological Features	220
<i>V.S. Abbiramy and A. Tamilarasi</i>	
Digital Video Watermarking Using Motion Detection and Singular Value Decomposition	229
<i>Sanjana Sinha, Swarnali Pramanick, Ankul Jagatramka, Prajnat Bardhan, Dipak K. Kole, and Aruna Chakraborty</i>	
Mining Association Rules Using Hash-Index Technique	239
<i>R.B. Geeta, Omkar Mamillapalli, Shasikumar G. Totad, and P.V.G.D. Prasad Redd</i>	
Iris Liveness Detection for Semi-transparent Contact Lens Spoofing	249
<i>Niladri B. Puhan, Sudha Natarajan, and A. Suhas Hegde</i>	
Stego-Image Generator (SIG) - Building Steganography Image Database	257
<i>P. Thiyagarajan, G. Aghila, and V. Prasanna Venkatesan</i>	
Review on OCR for Handwritten Indian Scripts Character Recognition	268
<i>Munish Kumar, M.K. Jindal, and R.K. Sharma</i>	
Satellite Image Processing Using Discrete Cosine Transform and Singular Value Decomposition	277
<i>B.K. Ashish, A. Kumar, and P.K. Padhy</i>	
Computer Science, Engineering and Information Technology	
Physical Layer Authentication in Wired and Wireless Communication Systems	291
<i>Chandra Mouli and Sanjay Singh</i>	
Cloud SaaS: Models and Transformation	305
<i>Ritu Sharma and Manu Sood</i>	
Future Robotics Memory Management	315
<i>S. Vijaykumar and S.G. Saravanakumar</i>	
Impact of Outliers on Anonymized Categorical Data	326
<i>K. Venkata Ramana, V. Valli Kumari, and K.V.S.V.N. Raju</i>	

Privacy in Mobile Ad Hoc Networks	336
<i>B. Jhansi Vazram, V. Valli Kumari, and J.V.R. Murthy</i>	
Prototype Based Performance Prediction of Web Services	346
<i>Ch. Ram Mohan Reddy, D. Evangelin Geetha, K.G. Srinivasa, T.V. Suresh Kumar, and K. Rajani Kanth</i>	
The Goal of Securing Mobile Ad-Hoc Network and Solutions	355
<i>Najiya Sultana and S.S. Sarangdevat</i>	
Elements of Cloud Computing: A Perspective on Service Oriented Enterprises (SOEs)	366
<i>R. Suchithra, R. Selvarani, and Dhinaharann Nagamalai</i>	
Extraction of Urban Road Network by Performing Edge – Aided Classification in Remote Sensing	378
<i>K. Rajalakshmi, A.L. Dhinakaran, D. Murugan, and S. Singaravelan</i>	
Analysis of Texture Patterns in Diagnosing Osteoporosis Using Proximal Femur X-ray Images	383
<i>V. Sapthagirivasan and M. Anburajan</i>	
Energy Efficient Data Aggregation in Cognitive Sensor Networks	393
<i>Sumalatha Ramachandran, Aswin Kumar Gopi, Giridara Varma Elumalai, and Murugesan Chellapa</i>	
VLSI Architecture for Compressed Domain Video Watermarking	405
<i>N. Mohankumar, M. Nirmala Devi, D. Badari Nath, and Arun Scaria</i>	
Decimal Floating Point Multiplier Using Double Digit Decimal Multiplication	417
<i>Rekha K. James, K. Poulose Jacob, and Sreela Sasi</i>	
Fuzzy Cost Based Multicast Routing for Mobile Ad-Hoc Networks with Improved QoS	429
<i>G. Santhi and Alamelu Nachiappan</i>	
Cooperative Resource Allocation for Primary and Secondary Users with Adjustable Priorities in Cognitive Radio Networks	438
<i>Behdis Eslamnour, S. Jagannathan, and Maciej Zawodniok</i>	
A Method for Identifying and Analyzing Authentication Mechanisms and a Case Study	449
<i>Sérgio Ribeiro and Danilo Suiama</i>	

An AI Based Framework Proposal for Real Time Robot Design Using Framsticks	460
<i>S. Raja Mohamed and P. Raviraj</i>	
Analysis and Comparision of Position Based Routing Protocol in Mobile Adhoc Network	468
<i>Vimal Upadhyay, Gori Shankar, Amritpal, and Jai Balwan</i>	
Author Index	477

Fuzzy Soft Thresholding Based Hybrid Denoising Model

Sudipta Roy¹, Nidul Sinha², and Asoke Kr. Sen³

¹ Department of IT, Assam University, Silchar - 788011, Assam
sudipta.it@gmail.com

² Electrical Engineering Department, NIT, Silchar - 788010, Assam
nidulsinha@hotmail.com

³ Department of Physics, Assam University, Silchar - 788011, Assam
asokesen@yahoo.com

Abstract. This paper proposes a denoising model hybridized using wavelet and bilateral filters with fuzzy soft thresholding. The parameters of the proposed model are optimized with floating point genetic algorithm (FPGA). The model optimized with one image is used as a general denoising model for other images like Lena, Fetus, Ultrasound, Xray, Baboon, and Zelda. The performance of the proposed model is evaluated in denoising images injected with noises in different degrees; moderate, high and very high, and the results obtained are compared with those obtained with similar hybrid model with wavelet soft thresholding. Results demonstrate that the performance of the proposed model in terms of PSNR and IQI in denoising most of the images is far better than those with similar model with wavelet soft thresholding. It has also been observed that the hybrid model with wavelet soft thresholding fails to denoise images with very high degree of noises while the proposed model can still be capable of denoising.

Keywords: Image denoising, fuzzy soft thresholding, wavelet thresholding, bilateral filter.

1 Introduction

Denoising of images is an indispensable task in image processing. Lot of works has been reported in the development of efficient denoising models. Thresholding is the key idea in many denoising models especially with wavelet based models. But, the choice of an appropriate value for thresholding is the major challenge. Because, the value of the threshold decides what coefficients should be shedded or shrunked and to what extent so that noise is eliminated while image information is not lost. But, there is no linear relationship between noise contents and image contents in spatial domain in an image. As such, it becomes very difficult to accurately and precisely determine the threshold value or the level of shrinkage.

Many denoising methods have been proposed over the years, such as the Wiener filter [1], [2], wavelet thresholding [3], anisotropic filtering [4], bilateral filtering [5], total variation method [6], and non-local methods [7]. Among these methods, wavelet thresholding has been reported to be a highly successful method. In wavelet

thresholding, a signal is decomposed into approximation (low-frequency) and detail (high-frequency) subbands, and the coefficients in the detail subbands are processed via hard or soft thresholding [2], [8], [9], [10]. The hard thresholding eliminates (sets to zero) coefficients that are smaller than a threshold as discussed earlier; the soft thresholding shrinks the coefficients that are larger than the threshold as well. The main task of the wavelet thresholding is the selection of the threshold and the effect of denoising depends on the selected threshold: a bigger threshold will throw off the useful information and the noise components at the same time while a smaller threshold cannot eliminate the noise effectively. Donoho [3] gave a general estimation method for threshold, but the best threshold cannot be found by this method. Chang et al. [11] have used predictive models to estimate the threshold. It is a spatially adaptive threshold based on context modeling. They also presented data-driven threshold for image denoising in a Bayesian framework [12]. In the SURE Shrink approach [10], the optimal threshold value based on the Stein's Unbiased Estimator for Risk (SURE) is estimated. A major strength of the wavelet thresholding is the ability to treat different frequency components of an image separately; this is important, because noise in real scenarios may be frequency dependent. But, in wavelet thresholding the problem experienced is generally smoothening of edges.

The bilateral filter was proposed in [13] as an alternative to wavelet thresholding. It applies spatially weighted averaging without smoothing edges. This is achieved by combining two Gaussian filters; one filter works in spatial domain, the other filter works in intensity domain. Therefore, not only the spatial distance but also the intensity distance is important for the determination of weights [14]. Hence, these types of filters can remove the noise in an image while retaining the images. However, the filter may not be very efficient in removing any noise in the texture part of the image.

A wisely hybridized model designed using both wavelet thresholding and bilateral filtering reported in [15], which exploits the potential features of both wavelet thresholding and bilateral filter at the same time their limitations are overcome. To improve further the efficiency of the model, the concept of Fuzzy Logic (FL) is introduced in this work for finding a rational value for the threshold value under highly non-linear noisy conditions. As the task of finding the optimum values for the parameters of the hybrid filters obtained by trial and error method is time consuming and in addition, the values may not be the optimum as the parameters are not linearly related, Floating point Genetic Algorithm (FPGA), which is reported to be very efficient in finding optimal solutions for very large and non-linear optimization problems, is used to optimize the values of the parameters of the hybrid model proposed including the parameters of the fuzzy soft thresholding. Further, it is also investigated whether the model optimized for denoising a particular image is well capable of denoising other images. Also, the capability of the similar hybrid models; one with wavelet soft thresholding (WST) and the other with fuzzy soft thresholding (FST), is tested by injecting noise in different degrees till any one fails.

In view of the above, the objectives of the present works are:

- (i) To optimize the values of the parameters of the best hybrid model as reported in [15] using FPGA with WST for the image Lena and obtain the performance of the same optimized model for denoising highly noisy Lena and other images.

- (ii) To optimize the values of the parameters of the same hybrid model as in step (i) with FST using FPGA for the image Lena and obtain the performance of the same optimized model for denoising Lena and other images.
- (iii) And compare the results obtained with those in steps (i) and (ii) for different degree of noises.

Section 2 of the paper introduces the concept of Fuzzy Soft Thresholding and works on it. Performance criteria are discussed in section 3. Section 4 describes the proposed FL based model and results are discussed in section 5. Finally the conclusion is drawn in section 6.

2 FL Based Thresholding

FL Based Threshold Function:

From the fuzzy logic point of view [16], the operation of hard-thresholding in the DWT domain performs a zero-one membership function, which maps wavelet coefficients into two membership values of 0 (excluded as member in signal reconstruction) or 1 (included as member in signal reconstruction). Since the distinction between a noise generated wavelet coefficient and a signal generated wavelet coefficient according to their magnitude value is not a crisp one, there is a dilemma in setting an appropriate threshold value, and a small variation in the threshold value may result in dramatic change in the reconstructed signal due to the abrupt change in boundary. To enable each wavelet coefficient to contribute in signal reconstruction according to the significance of their magnitude values, a membership function is required to map each wavelet coefficient to a corresponding membership value between 0 and 1. For detail theory [16] may be referred.

The membership function considered for fuzzy soft thresholding in this work is as given in (1).

$$\mu_F(r_{m,n}) = -\frac{\mu(r_{m,n}) - \max(\mu(r_{m,n}))}{\max(\mu(r_{m,n}))} \quad (1)$$

where $\mu(r_{m,n}) = \frac{b}{b + e^{a(r_{m,n})^4} + e^{-a(r_{m,n})^4}}$ and where a and b are constants for controlling the width of the passband, transition band and stopband of the membership function.

The values of a and b are to be chosen wisely as increase in the value of a results in a narrower transition band (steeper slope), a narrower stopband and a wider passband, with the increase in the passband being larger than the decrease in the stopband while decrease in the value of b also results in a narrower transition band, a narrower stopband and a wider passband, but with the increase in the passband being smaller than the decrease in the stopband.

The optimum values of the parameters a and b for the proposed threshold function are optimized using floating point genetic algorithm (FPGA) together with the other parameters of the best reported hybrid model in [15] to optimize the PSNR value. The image quality index (IQI) is also calculated simultaneously to quantify the image quality of the denoised image.

3 Measurement of Performance

To judge the performance of the denoising techniques *Mean Squared Error* (MSE) and *Peak Signal to Noise Ratio* (PSNR) [17] are the automatic choice for the researchers. But a better PSNR does not imply that the visual quality of the image is good. To overcome this problem Image Quality Index (IQI) [18] is considered in this work as the second parameter for judging the quality of denoised images. Although the index is mathematically defined and does not explicitly employ the human visual system model, experiments on various image distortion types show that it exhibits surprising consistency with subjective quality measurement.

4 Proposed Model with Fuzzy Soft Thresholding Function

The model considered for drawing comparison with the proposed model in this paper is the best hybrid model reported in [15]. The model is illustrated in Fig. 1. The proposed model in this paper is the model with fuzzy soft thresholding (FST) function instead of wavelet soft thresholding (WST) function as shown in Fig. 1.

In the proposed model the input image is first denoised with the help of bilateral filter. Then the image is decomposed using wavelet based decomposition technique into four subbands and the Fuzzy soft thresholding technique is applied on all of the subbands to chopped off the noisy signals. After chopping off, the subbands are again combined to reconstruct the image. The parameters of the model including fuzzy soft thresholding function are optimized using FPGA.

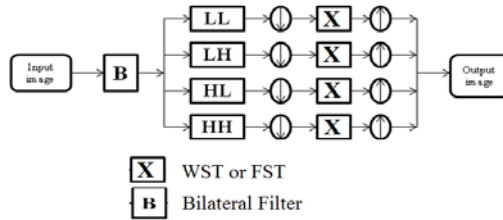


Fig. 1. Hybrid denoising model

5 Results and Discussions

The programmes are written in Matlab programming environment and model is tested on different types of images after injection of Gaussian noise in different levels; moderate, high and very high.

Case – 1:

The parameters of the considered model with WST are optimized using FPGA for image the Lena and the performance of the same, so optimized, in terms of PSNR and

IQI, in denoising other types of images like Fetus, Ultrasound (ultrasound images), Xray (x-ray image), Baboon, and Zelda (natural images) including Lena with different degree of noise is reported in table 1.

Case -2:

The parameters of the same hybrid model as in case -1 with FST are optimized using FPGA for the image Lena and the performance of the optimized model in terms of PSNR and IQI, in denoising Lena and other images as in case-1 is reported in table 2.

Comparison in the results for cases 1 & 2 are depicted in figs. 2 to 7.

Investigation of the results reveals that the performance of the hybrid model with FST in terms of PSNR and IQI is far better than that with WST. However, the IQI obtained with WST for image Baboon is better than that with FST. Hence, the proposed model with FST optimized for one image can be used as a general denoising model. Also, it is observed that the model with WST fails to denoise images injected with very high noises but the proposed model do it well.

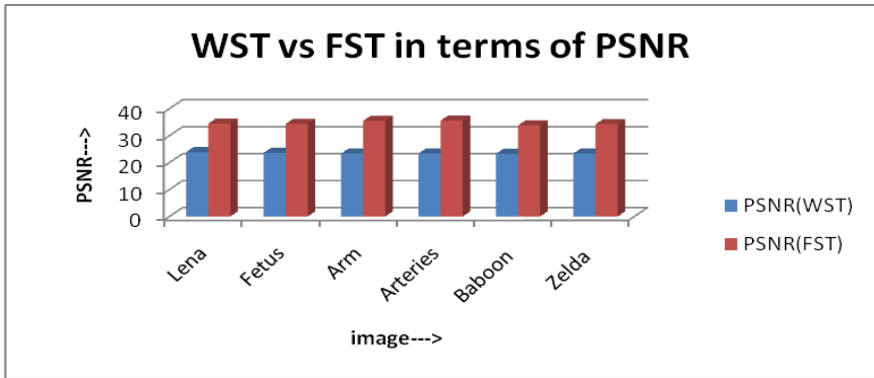


Fig. 2. Graphical Comparison of Table 1 & 2 in terms of PSNR

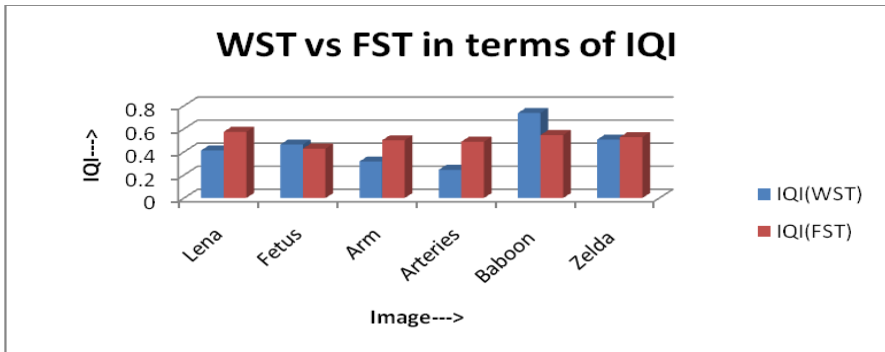


Fig. 3. Graphical Comparison of Table 1 & 2 in terms of IQI

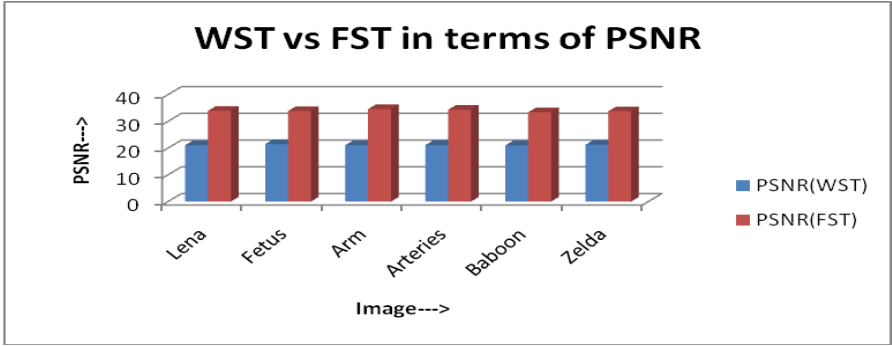


Fig. 4. Graphical Comparison of Table 3 & 4 in terms of PSNR

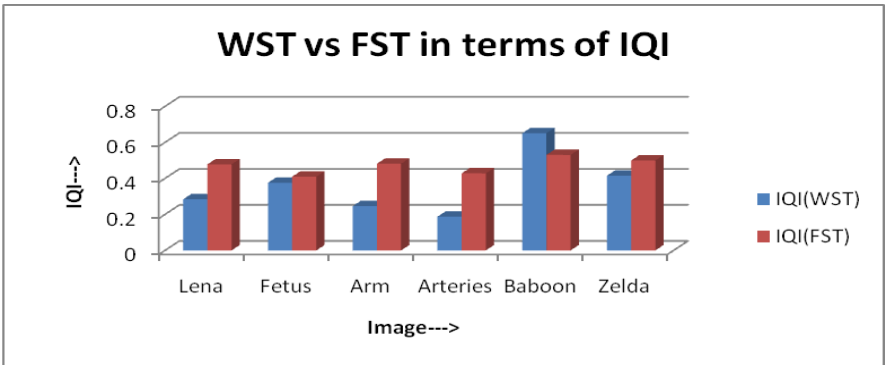


Fig. 5. Graphical Comparison of Table 3 & 4 in terms of IQI

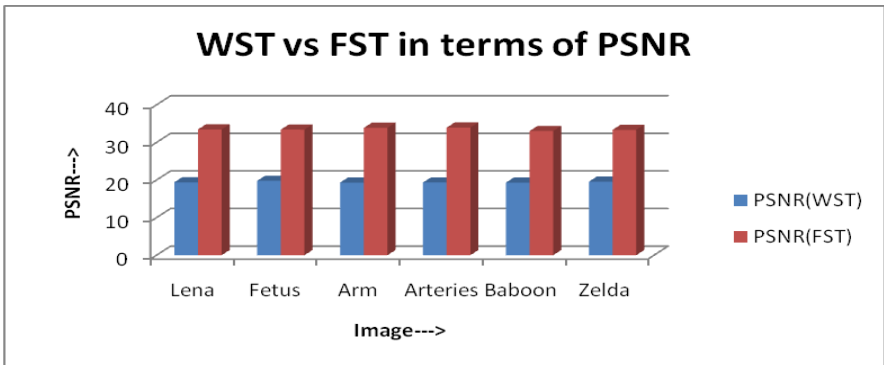


Fig. 6. Graphical Comparison of Table 5 & 6 in terms of PSNR

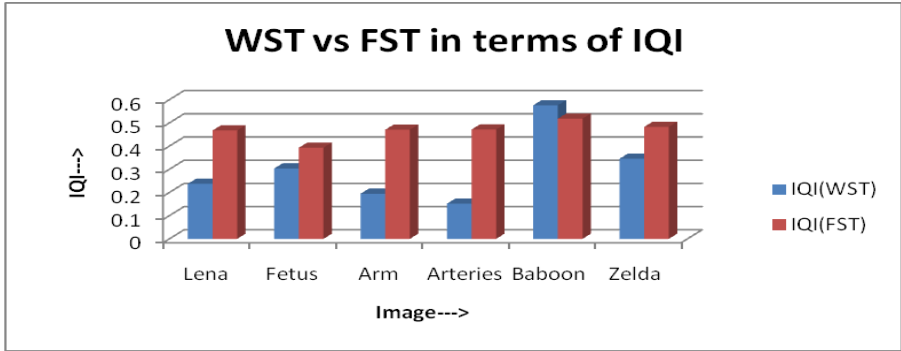


Fig. 7. Graphical Comparison of Table 5 & 6 in terms of IQI

The quality of the images after the denoising can be visualized from some of the prominent cases as reported below:

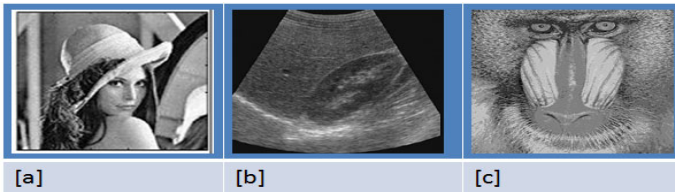


Fig. 8. Original images of [a] Lena, [b] Fetus and [c] Baboon

Noise Level	Noisy Image	Denoised (FST)	Denoised (WST)
Moderate			
High			
Very High			

Fig. 9. Lena: Noisy images and denoised images using FST and DST

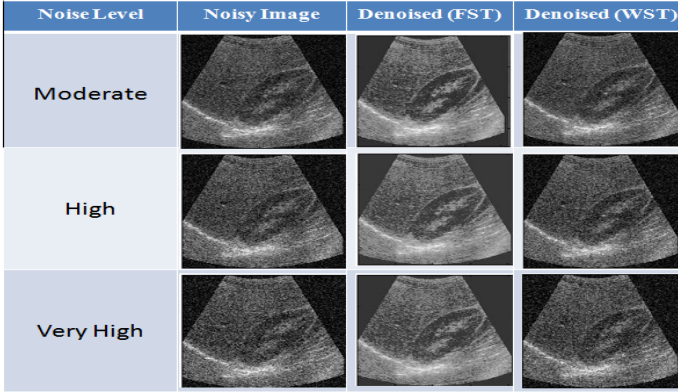


Fig. 10. Fetus: Noisy images and denoised images using FST and DST

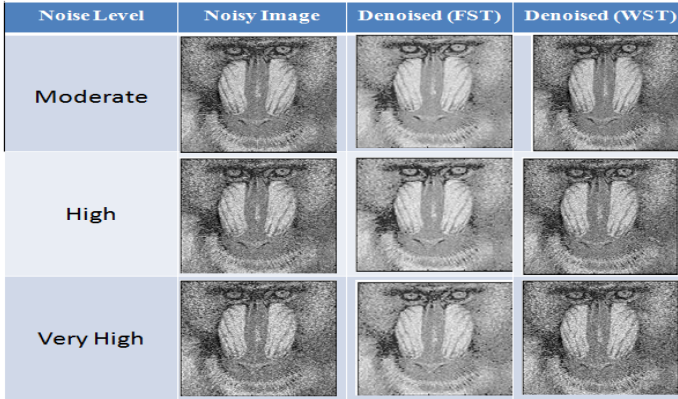


Fig. 11. Baboon: Noisy images and denoised images using FST and DST

Table 1. Best evaluated results of the proposed model with WST

Noise Level	Image	Lena	Fetus	Arm	Arteries	Baboon	Zelda
Moderate	PSNR	23.8870	23.6173	23.3411	23.3815	23.2674	23.4032
	IQI	0.4123	0.4633	0.3168	0.2448	0.7372	0.5062
High	PSNR	21.1001	21.4988	21.0982	21.1207	21.0369	21.2538
	IQI	0.2838	0.3758	0.2459	0.1874	0.6510	0.4151
Very High	PSNR	19.3884	19.7887	19.3240	19.3351	19.3068	19.5466
	IQI	0.2374	0.3035	0.1946	0.1515	0.5758	0.3455

Table 2. Best evaluated results of the proposed model with FST

Noise Level	Image	Lena	Fetus	Arm	Arteries	Baboon	Zelda
Moderate	PSNR	34.4547	34.3761	35.4929	35.5935	33.8136	34.2585
	IQI	0.5738	0.4282	0.5008	0.4884	0.5460	0.5276
High	PSNR	33.9736	33.8959	34.6231	34.3761	33.4333	33.7985
	IQI	0.4782	0.4097	0.4816	0.4282	0.5304	0.4996
Very High	PSNR	33.4605	33.4161	33.8864	33.9469	33.0440	33.3364
	IQI	0.4673	0.3930	0.4704	0.4713	0.5183	0.4828

6 Conclusion

The hybrid model with fuzzy soft thresholding (FST) is implemented in Matlab. The parameters of the model are optimised using FPGA. The performance of the proposed model optimized with image Lena in denoising other category of images is almost same as compared that if the model is optimized for individual category of images. The performance of the proposed model in denoising most of the categories of images is far better than that with similar model with wavelet thresholding. There is no significant improvement in performance for both the proposed model with fuzzy soft thresholding and the similar model with wavelet soft thresholding even if the models are optimized for each category of images separately. Hence, the proposed model with fuzzy soft thresholding optimized for one category of images can be recommended for denoising highly noisy images with good perceptual quality for most of the images.

References

1. Albert, C.T., Moore, J.R., Glaser, S.D.: Wavelet Denoising Techniques with Applications to Experimental Geophysical Data. *Signal Processing* 89(2), 144–160 (2009)
2. Shui, P.-l., Zhao, Y.-B.: Image Denoising Algorithm using Doubly Local Wiener Filtering with Block-adaptive Windows in Wavelet Domain. *Signal Processing* 87(7), 1721–1734 (2007)
3. Donoho, D.L., Johnstone, I.M.: Ideal spatial adaptation by wavelet shrinkage. *Biometrika* 81(3), 425–455 (1994)
4. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Analysis and Machine Intelligence* 12(7), 629–639 (1990)
5. Tomasi, C., Manduchi, R.: Bilateral filtering for gray and color images. In: *Proc. Int. Conf. Computer Vision*, pp. 839–846 (1998)
6. Rudin, L.I., Osher, S., Fatemi, E.: Onlinear total variation based noise removal algorithms. *Physica D* 60(1-4), 259–268 (1992)
7. Buades, Coll, B., Morel, J.: Neighborhood filters and PDE's. *Numerische Mathematik* 105(1), 1–34 (2006)
8. Donoho, D.L.: De-noising by soft thresholding. *IEEE Trans. on Inform. Theory* 41(3), 613–627 (1995)

9. Donoho, D.L., Johnstone, I.M.: Adapting to unknown smoothness via wavelet shrinkage. *Journal of the American Statistical Association* 90(432), 1200–1224 (1995)
10. Donoho, D.L., Johnstone, I.M., Kerkyacharian, G., Picard, D.: Wavelet shrinkage: Asymptopia. *Journal of Royal Statistics Society, Series B* 57(2), 301–369 (1995)
11. Chang, S.G., et al.: Adaptive wavelet thresholding for image denoising and compression. *IEEE Transactions on Image Processing* 9, 1532–1546 (2000)
12. Chang, S.G., Yu, B., Vetterli, M.: Spatially adaptive wavelet thresholding with context modeling for image denoising. In: *Proc. ICIP*, pp. 535–539 (1998)
13. Tomasi, C., Manduchi, R.: Lateral filtering for gray and color images. In: *Proc. Int. Conf. Computer Vision*, pp. 839–846 (1998)
14. Zhang, M., Gunturk, B.: A New Image Denoising Method based on the Bilateral Filter. In: *ICASSP*, pp. 929–932. IEEE, Los Alamitos (2008)
15. Roy, S., Sinha, N., Sen, A.K.: Performance Analysis of Wavelet and Bilateral Filter based Denoising Models: Optimized by Genetic Algorithm. In: *Proceedings of International Conference on Computing and Systems (ICCS) – 2010*, pp. 257–262 (2010)
16. Mendel, J.M.: Fuzzy logic systems for engineering: a tutorial. *Proceedings of IEEE* 83(3), 345–377 (1995)
17. Roy, S., Sen, A.K., Sinha, N.: VQ-DCT based Image Compression: A New Hybrid Approach. *Assam University Journal of Science and Technology* 5(II), 73–80 (2010)
18. Wang, Z., Bovik, A.C.: A Universal Image Quality Index. *IEEE Signal Processing Letters* 9(3) (2002)

Content-Based Image Retrieval Using a Quick SVM-Binary Decision Tree – QSVMBDT

I. Felci Rajam¹ and S. Valli²

¹ Department of M.C.A, St. Joseph's College of Engineering Jeppiar Nagar,
Old Mamallapuram Road, Chennai – 600119 & Research Scholar,
Department of C.S.E, Anna University, Chennai

² Department of CSE, College of Engineering, Guindy, Anna University Chennai
Sardar Patel Road, Guindy, Chennai - 600025, Tamil Nadu, India

Abstract. A Content-based image retrieval (CBIR) framework, which provides a quick retrieval using the SVM-Binary Decision Tree with prefiltering called as Quick SVM-BDT, is proposed. The Support Vector Machine-Pair Wise Coupling (SVM-PWC) and Fuzzy C-Mean (FCM) clustering are the supervised and unsupervised learning techniques respectively, used for the multi-class classification of images. In this system, the SVM based binary decision tree (SVM-BDT) is constructed for semantic learning, and it finds the semantic category of the query image. Similarity matching is done between the query image and only the set of images belonging to the semantic category of the query image. This reduces the search space. The search space is still reduced by prefiltering the images of that particular category which provides a very quick retrieval. Experiments were conducted using the COREL dataset consisting of 1000 images of 10 different semantic categories. The obtained results demonstrate the effectiveness of the proposed framework.

Keywords: Statistical Similarity Matching, SVM-Binary Decision Tree, Prefiltering, Clustering, Fuzzy C-Mean clustering, SVM-Pair Wise Coupling.

1 Introduction

The applications of a digital image in various domains are rapidly increasing. Different techniques have been proposed to increase the efficiency of image retrieval. Low level features, such as color, texture, edge and shape are extracted from the image. As the dimension of the feature vector extracted for the DB images is large, it occupies an enormous logical space in the DB, and increases the computational complexity of image retrieval. To reduce the dimension of the feature vectors, the proposed CBIR framework uses the PCA technique [13]. Sometimes, the images retrieved have features similar to those of the query image, but they vary in terms of their semantic interpretation, perceived by the user. This is known as the *Semantic gap* issue. To bridge the semantic gap, the DB images are filtered, based on their semantic category, using machine learning techniques. The supervised learning technique, SVM, can be used to solve multiclass problems [2] and [3]. The authors train the SVM using the DB images of known categories.

The Fuzzy C-Mean (FCM) clustering technique is an unsupervised learning technique, where the class labels are not available for training. When the query image is given, the FCM finds the cluster to which the query image belongs, and the DB images of that cluster alone are taken for comparison [2].

The multi class SVM-PWC uses $k(k-1)/2$ SVM binary classifiers to predict the class label of the image. Both the computation time and the number of classifiers required, are high. An efficient classification scheme, the SVM-BDT requires only $(k-1)$ SVM binary classifiers, and it avoids misclassification considerably. The authors [8] have used the SVM-BDT for solving multi-class problems. The training and testing time are reduced by the use of the SVM-BDT. They have applied this methodology for hand written character and digit recognition. This methodology is used for the CBIR framework in this proposed method.

Song Liu et al., [12] have developed an adaptive hierarchical multi-class SVM classifier for texture-based image classification. In this paper, they have presented a new classification scheme based on Support Vector Machines (SVM), and a new texture feature called texture correlogram for high-level image classification. To evaluate the performance of their system, they have used the COREL database of 900 images covering 9 different semantic categories. They conclude that the hierarchical classification scheme and texture features are effective for high-level image classification tasks, and that the classification scheme is more efficient than the other schemes while achieving almost the same classification accuracy. Here, they have considered only the texture feature of the images and not the other low level features such as color, shape and edge features.

In this QSVMBDT, the class label of the query image is predicted using the SVM-BDT; then, the probability estimates obtained from the SVM-PWC and the cluster membership values obtained from the FCM are used for prefiltering the DB images of the specific category. Therefore, a reduction has been achieved in search space and computation time.

The main contributions of this paper are as follows: For minimizing misclassification and reducing the computation time, the SVM based binary decision tree architecture is used. To reduce the search space for image retrieval, the SVM based binary decision tree architecture with prefiltering is proposed by combining the probability estimates of the SVM-PWC and the cluster membership values from FCM clustering.

The remainder of this section is organized as follows: Section 2 covers the implemented system architecture. Section 3 gives the experimental results and section 4 concludes the work.

2 System Architecture

The main modules of this system are feature extraction, dimensionality reduction by the PCA, class prediction using the SVM-BDT, prefiltering and statistical similarity measures. The architecture of the QSVMBDT system is shown in Fig. 1.

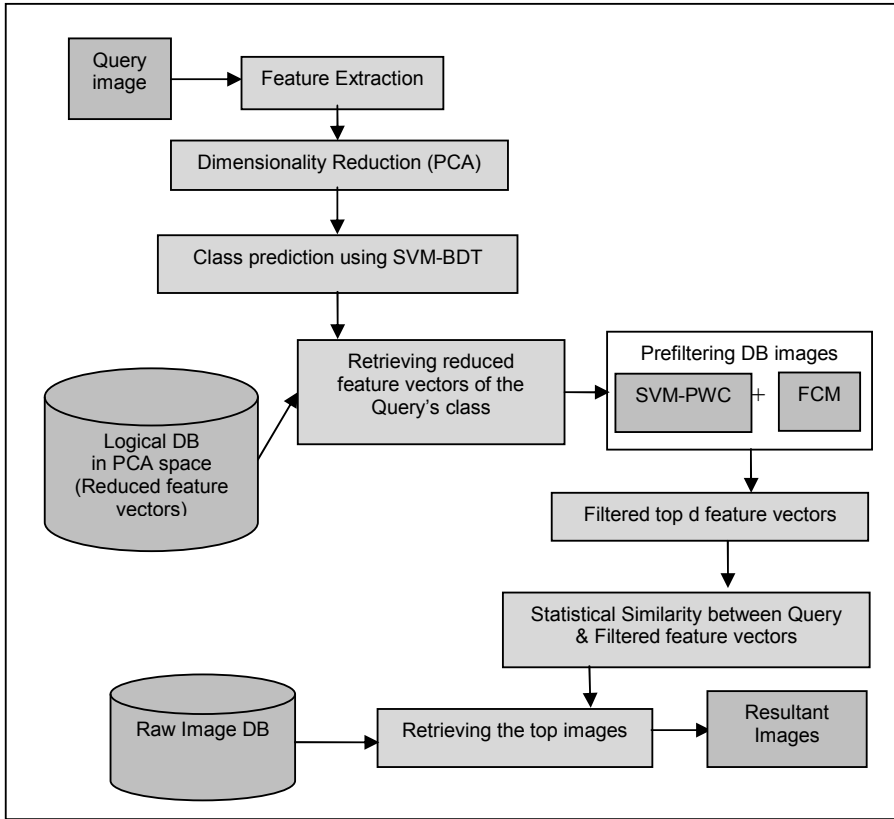


Fig. 1. System Architecture of the QSVMBDT

2.1 Feature Extraction

Low level features such as color, texture, edge and shape features are obtained for the query and DB images. As one feature vector does not work well, a combined feature vector is used as input for the classification. The first, second and third central moments of each color channel are used in representing the color feature vector in the Hue, Saturation and Value in HSV color space [14]. The nine dimensional $(\mu_H, \mu_S, \mu_V, \sigma_H, \sigma_S, \sigma_V, \gamma_H, \gamma_S, \gamma_V)$ color feature vector (f_c) is extracted. μ represents the mean, σ corresponds to the standard deviation and γ corresponds to the skewness of each color channel in the HSV [2].

The texture information is extracted from the gray-level co-occurrence matrix. A gray level co-occurrence matrix is defined as a sample of the joint probability density of the gray levels of two pixels separated by a given displacement d and angle θ [2]. The four co-occurrence matrices of four different orientations (horizontal 0° , vertical 90° and two diagonals 45° and 135°) are constructed. The co-occurrence matrix

reveals certain properties about the spatial distribution of the gray levels in the image. Higher order features such as energy, contrast, homogeneity, correlation, entropy are measured on each gray level co-occurrence matrix, to form a five dimensional feature vector. Finally, a twenty dimensional texture feature vector (f_t) is obtained by concatenating the feature vectors of each co-occurrence matrix [1].

The edge information contained in the images is obtained using canny edge detection (with $\sigma=1$, Gaussian masks of size =9, low threshold=1, and high threshold=255) algorithm. The corresponding edge directions are quantized into 72 bins of 5° each. The histograms are normalized with respect to the number of edge points in the image. Hence, a 72 dimensional edge feature vector (f_e) is obtained [2]. The shape feature vector of the image is obtained by calculating seven invariant moments [7]. These features are invariant under rotation, scale, translation and reflection of images. The central moments are calculated and a 7-dimensional shape vector (f_s) is obtained.

Let f_c, f_t, f_e, f_s be the feature vectors of the color, texture, edge and shape respectively. Then $f_{combined} = (f_c + f_t + f_e + f_s)$. Let \mathfrak{R}^d be the feature dimension of $f_{combined}$, where $d = (9+20+72+7=108)$. To reduce the logical size and high computational complexity, the PCA technique has been employed in the proposed system for reducing the dimension.

2.2 Principal Component Analysis

The PCA is the most powerful technique for the dimension reduction of the feature vector, as it reduces the dimension without much loss of information [13]. The dimension reduction corresponds to selecting the most representative feature attributes. The basic idea of the PCA is to find k linearly transformed components implying the maximum amount of variances in the input data [2]. The original feature vector in the \mathfrak{R}^d dimension for the query and database images are projected into the \mathfrak{R}^l dimension, where $l \ll d$. The projected feature vectors are used for SVM training using the binary decision tree.

2.3 Class Prediction Using the SVM-Binary Decision Tree

The SVM-BDT solves the multi class pattern recognition problem using a binary tree, in which each node makes a binary decision using the SVM binary classifier [8], [9].

The reduced low-level feature vector $X_i \in \mathfrak{R}^l$ for the DB images is used in training the SVM, and for predicting the class label of the query image. As it reaches the leaf node, labels are predicted for the query image, using only the (K-1) classifiers. The gravity center for each class of images is found by taking the mean of the reduced feature vectors of each category of the DB images [4], [8]. The Euclidean distance is used to create a distance matrix of dimension $K \times K$ where K is the number of classes. Its cells are filled by calculating the Euclidean distance between the gravity centers of the i^{th} and j^{th} classes ($1 \leq i, j \leq K$).

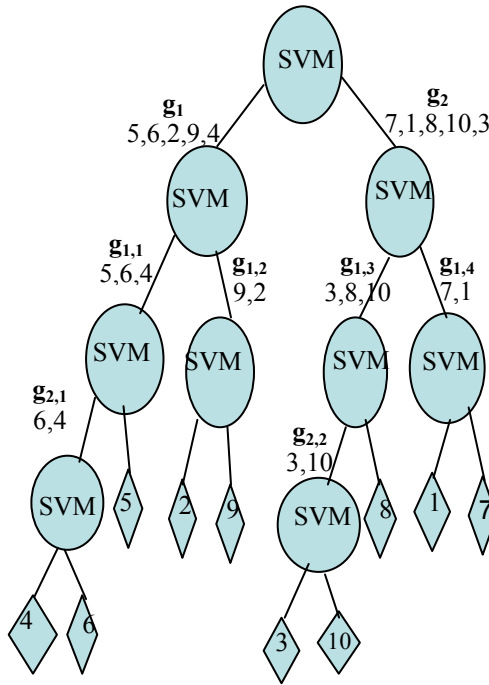


Fig. 2. SVM based binary decision tree

The two classes that have the largest Euclidean distance are assigned to each of the two clustering groups, and their gravity centers are taken as the cluster centers [8], [9]. After this, a pair of classes that is closest to the clustering groups is found, and assigned to the corresponding group. The cluster center of each clustering group is modified as the gravity center of the newly added class of that group. This process continues until all the classes are assigned to one of the two possible groups of classes. Then each group is divided into two sub groups, until there is only one class per group [8], [9]. This procedure leads to a binary tree for the SVM-BDT. Fig. 2 illustrates the grouping of the 10 classes. After finding the Euclidean distance, classes c_5 and c_7 are the farthest, and are assigned to the group g_1 and g_2 respectively. Closest to group g_1 is class c_6 and closest to group g_2 is c_1 . Hence, c_6 is assigned to group g_1 and c_1 is assigned to group g_2 . In the next step, c_2 is assigned to group g_1 and c_8 is assigned to group g_2 . In the next step, c_9 is assigned to group g_1 and c_{10} is assigned to group g_2 . In the next step, c_4 is assigned to group g_1 and c_3 to group g_2 . This completes the first round of grouping that defines the classes that will be transferred to the left and right sub tree of the node. The SVM classifier in the root is trained by considering the samples from the classes $\{c_5, c_6, c_2, c_9, c_4\}$ as positive and samples from the classes $\{c_7, c_1, c_8, c_{10}, c_3\}$ as negative.

The grouping procedure is repeated independently for the classes of the left and the right sub tree of the root, which results in grouping $\{c_5, c_6 \text{ and } c_4\}$ in $g_{1,1}$ and $\{c_9, c_2\}$ in $g_{1,2}$ on the left side of the sub tree, $\{c_3, c_8 \text{ and } c_{10}\}$ in $g_{1,3}$ and $\{c_7, c_1\}$ in $g_{1,4}$ on the right side of the sub tree. At the next level, $\{c_6 \text{ and } c_4\}$ is grouped in $g_{2,1}$ on the left side of the sub tree and $\{c_3 \text{ and } c_{10}\}$ in $g_{2,2}$ on the right side of the sub tree.

After building the SVM-BDT, a binary class SVM classifier is placed at each non-leaf node to train the classes that are on the left side as positive samples, and classes on the right side to be negative samples [8]. The reduced features of the query image are given to the binary classifier at the root level. According to its result, it branches the query image either to the left or to the right side of the root, and the process continues till it reaches the leaf level. At the leaf level, the class label of the query image is obtained. The distance is calculated between the query image and the DB images of the particular class. The distances are sorted in the ascending order and the top 15 images are displayed as the result. The Euclidean, Bhattacharya and Mahalanobis distances are used and given in equations (1) to (3) where q is the query image and t is the target image [2], [10].

$$D_{\text{Euc}}(q_i, t_i) = \sum_{i=1}^n (q_i - t_i)^2 \quad (1)$$

$$D_{\text{Bhat}}(q, t) = \frac{1}{8} (\mu_q - \mu_t)^T \left[\frac{\Sigma_q - \Sigma_t}{2} \right]^{-1} (\mu_q - \mu_t) + \frac{1}{2} \ln \frac{|\Sigma_q + \Sigma_t|/2}{\sqrt{|\Sigma_q| |\Sigma_t|}} \quad (2)$$

$$D_{\text{Maha}}(q, t) = (\mu_q - \mu_t)^T \Sigma^{-1} (\mu_q - \mu_t) \quad (3)$$

where μ is the mean and Σ is the covariance matrix.

2.4 Multi Class Classification Using SVM-Pair Wise Coupling

Multi class classification means assigning a class from one of the k predefined classes for each image under consideration. A variety of techniques are available for the decomposition of the multi-class problem into several binary class problems, using Support Vector Machines as binary classifiers [3]. This is a supervised learning technique, where the class labels are already available for the database images.

The one-against-one technique also known as the ‘‘pairwise coupling’’ (PWC) technique is used to predict the class label for the query image. This method trains $K(K-1)/2$ binary classifiers for k mutually exclusive classes, each of which provides a partial decision for classifying a data point. It then combines the output of all the binary classifiers to form a class prediction. For the query image, each one of the $K(K-1)/2$ binary classifiers votes for one class. The class label of the query image is the class for which the maximum vote occurs [3].

The output of the SVM-PWC is the label of the image along with its probabilities regarding each K category. The image belongs to the class for which the probability estimate is high [3].

<Image no> <probability estimate1> ... <Probability estimate K>

The probability estimates used for predicting the class labels are used for the prefiltering of images.

2.5 Fuzzy C-Mean Clustering

In the unsupervised machine learning technique, the class labels are not available as in the case of the supervised classification. The natural clusters in the image set are to be identified and clustered. The images within a cluster are similar with respect to certain aspects, and can be used for the categorization of images [6], [15].

The Fuzzy *c*-mean technique is the most widely used technique for clustering. The degree of membership of a data item to a cluster is between [0, 1] when the clusters are fuzzy.

For a given query image, the output of the FCM is the membership value of the image with each of the *K* classes. The query image belongs to the class for which the membership value is high [2].

<Image no> <membership value 1> ... <membership value K>

2.6 SVM Binary Decision Tree with Prefiltering of Images

Prefiltering of images is done, by using the probability estimates obtained from the SVM-PWC, and the cluster membership values obtained from the FCM. To reduce the DB search space and computation time, the DB images are prefiltered based on the label predicted using the SVM-BDT, and use that label to obtain the probability estimate of the class using the SVM-PWC, and the membership value of the class using the FCM. The implemented prefiltering algorithm is given below.

Prefiltering Algorithm

Step1: *N* represents the number of DB images and *S* represents the semantic classes. $S = \{S_1, S_2, \dots, S_K\}$.

Step2: The feature vectors of the DB images are extracted as $f_i = (f_c + f_t + f_e + f_s) \in \mathfrak{R}^d$ and the dimension of the feature vector is reduced by the PCA as $X_i \in \mathfrak{R}^l$, for $i=1, 2, \dots, N$ where $l \ll d$.

Step3: The gravity center of each category of images is found by considering the reduced feature vectors of each category. The SVM-BDT is constructed, using these gravity centers. The SVM binary classifier at each non-leaf node of the SVM-BDT is trained so as to predict the label of any query image.

- Step4:** The multi class SVM-PWC is trained with the subset of the DB images of K categories. For each DB image, the probability estimates for all K categories is found, which provides the probability of its membership to that particular category.
- Step5:** The FCM algorithm is used to find the cluster membership matrix U for the DB images, which provides the degree to which a particular image belongs to a particular cluster.
- Step6:** The feature vector of the query image is found $f_q = (f_c + f_t + f_e + f_s) \in \mathfrak{R}^d$ and it is projected from $f_q \in \mathfrak{R}^d$ into $X_q \in \mathfrak{R}^l$ by the PCA.
- Step7:** For the query image q , the reduced feature vector X_q is given to the SVM-BDT, and it predicts the class label s to which it belongs.
- Step8:** The probability estimates (φ_s) determined by the SVM-PWC and the cluster membership values (ϑ_s) found using the FCM, are taken for the DB images in the query image semantic class s .
- Step9:** The union of φ_s and ϑ_s is found as $q_i = \varphi_s + \vartheta_s$. It is normalized with respect to the total number of images in that class. The normalized values are sorted in the descending order and the top d images are taken as the prefiltered images.
- Step10:** The distance between the query image and the prefiltered images is found, and they are sorted in the ascending order. The top k images with the least distance are the resultant images.

Thus, the prefiltering algorithm reduces the search space, and thereby reduces the testing time, while maintaining the same accuracy rate.

3 Experimental Results

In order to verify the effectiveness and efficiency of the proposed system, experiments were conducted on the COREL dataset [16] consisting of 1000 images of sizes 256 x 384 and 384 x 256. A fully labeled DB is the training sample for the SVM. After close examination of the DB, 10 semantic categories, each consisting of one hundred images were selected. The different semantic categories consist of pictures of African faces, sea shore views, buildings, buses, dinosaurs, elephants, roses, horses, mountains and dishes.

The obtained feature vector dimensionality is reduced $\mathfrak{R}^d \rightarrow \mathfrak{R}^l | d = 108, l = 23$ by the PCA. The first twenty three eigen values related to the first twenty three high variances, which gives the cumulative variance of 50.5, and it is the 90% of the total variance, taken as the reduced feature vector [2].

The advantage of using the smaller set of eigen vectors is that it can increase the retrieval speed when large DBs are searched using statistical similarity measures. The principal component values are stored in the logical DB for later access.

For the SVM-BDT, the gravity centers are calculated by obtaining the mean for each reduced feature vector for the images in each semantic category. Thus ten gravity centers are calculated for the considered 10 semantic categories. Then the Euclidean distances are applied to the gravity centers, and further grouping occurs till it reaches the leaf node. The SVM binary classifier is used in training each level node, and it divides the group into two sub groups. Only nine SVM binary classifiers are needed to perform the multi class classification. The feature vector of the query image is reduced using the PCA, and given to the SVM-BDT classification for predicting the semantic class of the query image.

For SVM-based image classification, the radial basis kernel function (RBF) works well when the relation between the attributes is non-linear. The two parameters considered for the RBF kernel are C and σ . The kernel parameter σ controls the shape of the kernel and the regularization parameter C controls the tradeoffs between margin maximization and error minimization [5]. To calculate C and σ , a 10-fold cross validation (CV) is performed on the training set. Using the 10-fold cross validation, the training sets are divided into 10 subsets of equal size. Sequentially, one subset is tested against the remaining nine subsets. Thus, the whole training set is tested once, and the accuracy of the CV is the percentage of data correctly classified. The *LIBSVM* has been used for implementing the SVM-PWC [11].

The weighting components $m=2.0$, the termination criteria $\varepsilon=0.001$ and the number of clusters $c=10$ are set for the Fuzzy C-mean calculation, which gives the membership value for each of the DB images with respect to the number of clusters.

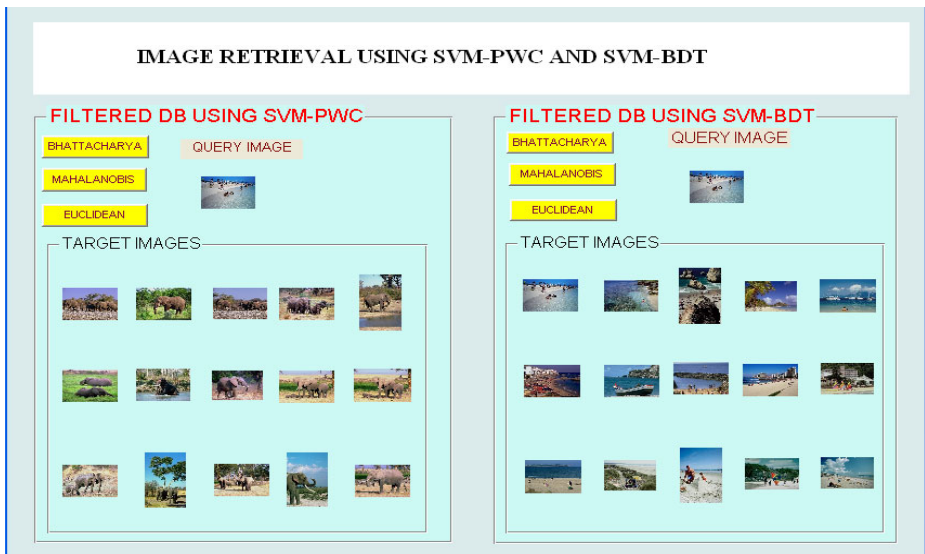


Fig. 3. Image Retrieval based on the SVM-PWC and SVM-BDT

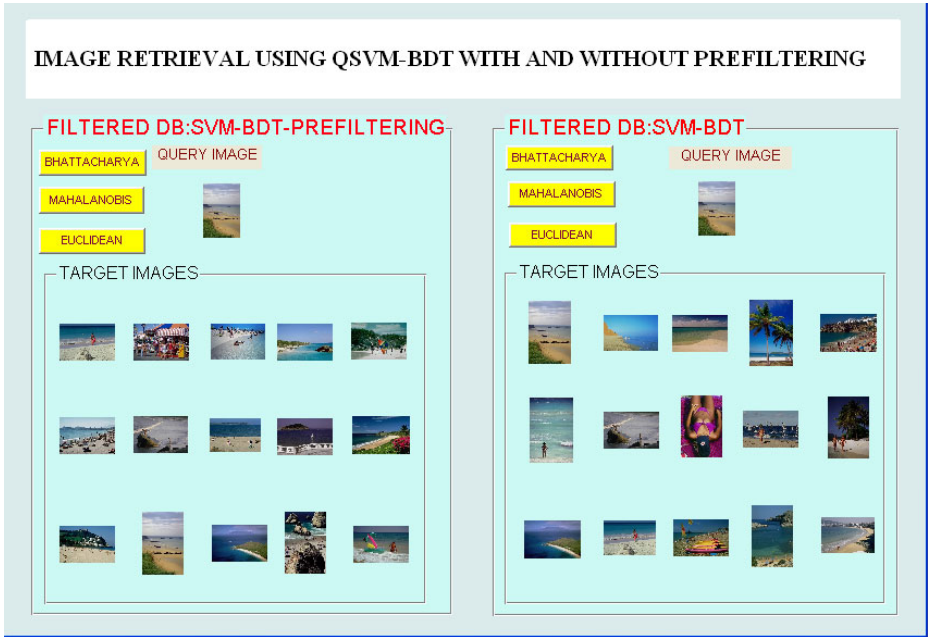


Fig. 4. Image Retrieval using the QSVM-BDT with prefiltering and SVM_BDT without Prefiltering

Experiments were done on the total DB and the filtered DB, with the SVM-PWC and SVM-BDT techniques. The comparison of the image retrieval using the SVM-PWC and SVM-BDT is shown in Fig. 3. The misclassification of the image shown in the SVM-PWC is avoided in the proposed SVM-BDT. In fig. 3, the query image is misclassified using the SVM-PWC; the correct classification of the 15 top images of the same category is obtained by the SVM-BDT. The snapshots of the image retrieval system for the filtered DB using the QSVM-BDT with prefiltering, and the SVM-BDT without prefiltering are shown in Fig. 4. It shows that the QSVM-BDT with prefiltering reduces the search space and testing time, while maintaining the accuracy of retrieval.

Table 1. Results of the Corel image data set

Measured in terms	Classifier		
	SVM-PWC	SVM-BDT	QSVMBDT with Prefiltering
Accuracy rate (%)	74.5	97.6	97.6
Training time(sec)	1. 36	1. 19	1.23
Testing time(sec)	1. 34	0. 05	0.03

The accuracy of the SVM-BDT is higher than that of the SVM-PWC. The QSVMBDT system is tested and the results are shown in Table 1. From the table, it is seen that the accuracy rate of the QSVMBDT technique with prefiltering, and that of the SVM-BDT technique without prefiltering, is far better than that of the SVM-PWC. The number of binary classifiers needed for the SVM-PWC is $K(K-1)/2$, but the number of binary classifiers needed for the SVM-BDT is only $(K-1)$. Since the number of classifiers is reduced, the testing time is reduced in the SVM-BDT. The testing time of the QSVMBDT with prefiltering is further reduced, because prefiltering reduces the search space still more. The training time of the QSVMBDT is slightly increased because of the prefiltering operation. Since the training time is a single time operation, it could be neglected. Fig. 5 shows the comparison graph of the three methods the SVM-PWC, SVM-BDT and QSVMBDT.

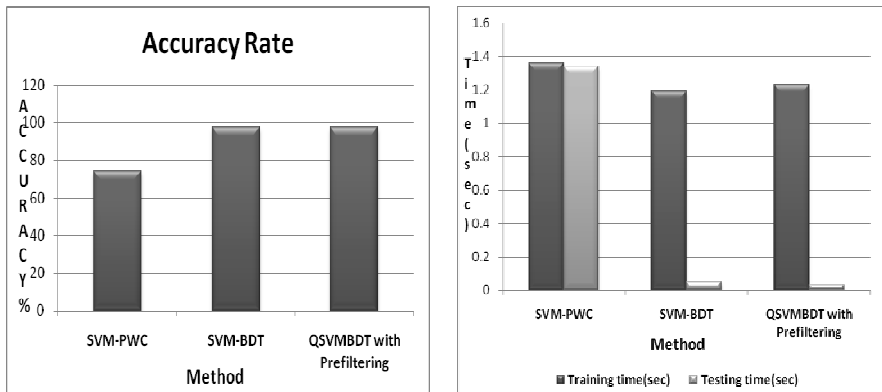


Fig. 5. Comparison graph of the Accuracy rate, training and testing time

4 Conclusion

This paper proposes a Quick SVMBDT framework, that performs an efficient and quick retrieval using the SVM based binary decision tree architecture and prefiltering. The low level features of the images are extracted, and the gravity center for each class of images is calculated, and the distance between the gravity centers is used in constructing the SVM-BDT. The SVM-BDT is used to predict the class label of the query image. Thus, only the images belonging to this class label are compared with the query image, which reduces the search space and thereby the retrieval time. The prefiltering technique reduces the search space further by combining the probability estimates of the SVM-PWC and the cluster membership values from the FCM clustering, and taking only the top images for comparison. Thus the SVMBDT technique for the CBIR framework exhibits a high accuracy rate, minimal training and testing time, compared to the earlier approach, the SVM-PWC, because the SVM-BDT approach requires only $(k-1)$ classifiers, while the SVM-PWC requires $K(K-1)/2$ classifiers. The QSVMBDT with prefiltering also depicts the same accuracy as the SVM-BDT, while reducing the testing time. Hence, the proposed CBIR framework using the QSVMBDT can be used as a front-end for image search, which yields high accuracy, and takes less retrieval time.

References

1. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. *IEEE Trans. Syst., Man, Cybern. SMC-3*(6), 610–621 (1973)
2. Rahman, M.M., Bhattacharya, P., Desai, B.C.: A framework for Medical Image Retrieval Using Machine Learning and statistical Similarity matching Techniques with Relevance Feedback. *IEEE Transactions on Information Technology in Biomedicine* 11(1), 58–69 (2007)
3. Wu, T.-F., Lin, C.-J., Weng, R.C.: Probability Estimates for Multiclass Classification by Pairwise Coupling. *Journal of Machine learning Research*, 975–1005 (August 2004)
4. Zhang, G., Jin, W.: Automatic Construction Algorithm for Multi-class Support Vector Machines with Binary Tree Architecture. *International Journal of Computer Science and Network Security* 6(2A), 119–126 (2006)
5. Chapelle, O., Haffner, P., Vapnik: SVMs for histogram-based image Classification. *IEEE Transaction on Neural Networks* 10(5), 1055–1064 (1999)
6. Yang, J.F., Hao, S.S., Chung, P.C.: Color Image Segmentation using Fuzzy C-Means and Eigenspace Projections. *Signal Process.* 82, 461–472 (2002)
7. Dudani, S.A., Breeding, K.J., McGhee, R.B.: Aircraft identification by moment invariants. *IEEE Trans. on Computers* C-26(1), 39–46 (1977)
8. Madzarov, G., Gjorgjevikj, D., Chorbev, I.: A Multi-class SVM Classifier Utilizing Binary Decision Tree. *Informatica* 33, 233–241 (2009)
9. Madzarov, G., Gjorgjevikj, D., Chorbev, I.: Multi-Class classification using support vector machines in binary tree architecture. In: *Int. Scientific Conf., Gabrovo* (2008)
10. Rahman, M. M., Bhattacharya, P., Desai, B.C.: Similarity Searching in Image Retrieval with Statistical Distance Measures and Supervised Learning. In: Singh, S., Singh, M., Apte, C., Perner, P. (eds.) *ICAPR 2005*. LNCS, vol. 3686, pp. 315–324. Springer, Heidelberg (2005)
11. Hsu, C.-W., Chang, C.-C., Lin, C.-J.: A Practical guide to Support Vector Classification, <http://www.csie.ntu.edu.tw/~cjlin>
12. Liu, S., Yi, H., Chia, L.-T., Rajan, D.: Adaptive Hierarchical Multi-class SVM Classifier for Texture-based Image Classification. In: *IEEE International Conference on Multimedia and Expo*. (2005)
13. Jain, A.k., Chandrasekaran, B.: Dimensionality and sample size consideration in pattern recognition practice. In: Krishnaiah, P.R., Kanal, L.N. (eds.) *Handbook of Statistics*, vol. 2, pp. 835–855. North Holland, Amsterdam (1987)
14. Stricker, M., Orengo, M.: “Similarity of color images: Storage and retrieval for image and video databases. In: *Proc. SPIE*, vol. 2420, pp. 381–392 (1995)
15. Bezdek, J.C., Pal, M.R., Keller, J., Krishnapuram, R.: *Fuzzy Models and Algorithms for pattern recognition and Image processing*. Kluwer Academic Publishers, Norwell (1999)
16. COREL image database, <http://wang.ist.psu.edu/docs/related/>

Implementation of LabVIEW Based Intelligent System for Speed Violated Vehicle Detection

G. Adinarayana¹, B. Lakshmi Sirisha², K. Sri Rama Krishna³,
and M. Kantikiran⁴

¹ PG Scholar, Department of ECE, V.R. Siddhartha Engineering College, Vijayawada,
Andhrapradesh, India

adinarayana18@yahoo.com

² Lecturer, Department of ECE, V.R. Siddhartha Engineering College, Vijayawada,
Andhrapradesh, India

suravarapuls@yahoo.co.in

³ Professor and HOD, Department of ECE, V.R. Siddhartha Engineering College, Vijayawada,
Andhrapradesh, India

srk_kalva@yahoo.com

⁴ Lecturer, Department of ECE, V.R. Siddhartha Engineering College, Vijayawada,
Andhrapradesh, India

kantikiran81@google.com

Abstract. This research intends to develop the speed violated vehicle detection system using image processing technique. Overall works are the software development of a system that requires a video scene, which consists of the following components: moving vehicle, starting reference point and ending reference point. A dedicated digital signal processing chip is used to exploit computationally inexpensive image-processing techniques over the video sequence captured from the fixed position video camera for estimating the speed of the moving vehicles. The moving vehicles are detected by analyzing the binary image sequences that are constructed from the captured frames by employing the interface difference or the background subtraction Algorithm. A novel adaptive Thresholding method is proposed to binarize the outputs from the interframe difference and the background subtraction techniques. The system is designed to detect the position of the moving vehicle in the scene and the position of the reference points and calculate the speed of each static image frame from the detected positions and report, to speed violated vehicle information to the authorised remote station.

Keywords: Background Subtraction, Interframe Difference, Image Processing, Adaptive Thresholding, Colour Modelling, Moving Object Detection, Tracking, and Vehicle Speed Measurement.

1 Introduction

Vehicle speed monitoring is important for enforcing speed limit laws. It also tells the traffic conditions of the monitored section of the road of interest. Speed camera

systems are widely used as a deterrent tool to prevent drivers from speeding. They mainly consist of three major units: 1) vehicle detectors; 2) a fixed-position video camera; and 3) a processor to control the overall system. The vehicle detectors are generally installed in pairs [1] with a known physical distance between them, and signals that originated in the existence of a vehicle are used to estimate the time that it takes for the corresponding vehicle to travel in between two detectors. The speed of the vehicle is then estimated according to the known distance and the estimated travel time between detectors. Then, the controlling processor triggers the camera to capture the image of the speeding vehicle according to the speed regulations. The vehicle detectors could be categorized into two types: 1) hardware-based detectors and 2) software-based detectors. The former is based on the electromagnetic principles and requires a dedicated hardware. The latter, which is adopted in this paper, performs vehicle detection using complex Image processing techniques over vehicle detectors can be categorized into four main groups: 1) inductive loop detector; 2) laser detector; 3) optical detector; and 4) weight detector.

An inductive loop detector is a wire loop that is embedded into the road's surface [1]. The wire loop emits a magnetic field that can detect vehicles by the metal contained within a vehicle and only detects vehicles that travel directly above them. Loop detectors are well known and well-studied technologies that have been widely used within many surveillance projects. Installation is relatively inexpensive, and power requirements are low. However, the main disadvantages related with them are that they need to be located within the pavement (hence subject to pavement damage) and installation and maintenance requires traffic disruption. For laser detector [2], a laser is located above the roadway, emitting a beam aimed at a photodiode array placed on the pavement. The vehicle detection is achieved when the vehicle breaks the laser beam and the photodiode array detects its presence. The laser detector can work in day and in night under the different weather conditions, but extreme temperatures could yield performance loss in their functionalities. The optical detector achieves vehicle detection using two light-activated optical sensors placed inside road studs located on the ground. Installation and maintenance of this technology is easy, but the technology is new, and conclusive results on accuracy are not yet available. The weight detector employs bending plates, piezoelectric sensors, or fiber-optic load sensors to detect vehicle weight.

Measuring weight is extremely useful for vehicle detection and classification and is rarely collected with other types of detector. The weight detectors have substantial error in measuring vehicle weight, and they are located within the pavement, which requires lane closures. For the software-based vehicle detectors, video cameras are placed alongside a roadway to achieve moving vehicle detection [4, 7]. The video cameras continuously acquire images of the traffic flow [3], which can be analyzed to extract a variety of information. Software-based vehicle detector is nonintrusive and does not require roadway disruption for installation or maintenance, but the performance of image analysis algorithms can suffer from poor weather, darkness, glare, or shadows [5]. Furthermore, high power computing processors are needed to perform computationally complex image processing algorithms in real time. Computationally inexpensive image processing techniques are adopted and developed for vehicle detection and tracking. Image processing is the technology, which is based on the software component that does not require the special hardware with a typical

video recording device and a normal computer [6]. We can create a speed detection device. By using the basic scientific velocity theory, we can calculate the speed of a moving vehicle in the video scene from the known distance and time, which the vehicle has moved beyond.

2 Binary Image Generation

The speed measurement is performed in binary image domain, i.e., each pixel is transformed into either “1” or “0” according to its motion information. To binarize the incoming input image and only detect the moving pixels, two different techniques are used: (1) interframe difference and (2) background subtraction.

2.1 Interframe Difference Technique

Let us assume that the input RGB image from the camera through a video card is represented as X_t^{RGB} of size $X_H \times X_W$ at time t . Furthermore, considering the region of interest, and let $I_t^{(RGB)}$ be the sub image of size $H \times W$ referring to the pixel intensity values of the spatial region labelled with R_0 . The difference image between consecutive frames at times t and $t - 1$ is computed by taking the absolute valued difference of gray scale representations of frames. The RGB image $I_t^{(RGB)}$ is converted to gray scale image I_t by using simple averaging of color channels, i.e.,

$$I_t = \frac{I_t^{(R)} + I_t^{(G)} + I_t^{(B)}}{3} \quad (1)$$

Where $I_t^{(R)}$, $I_t^{(G)}$, $I_t^{(B)}$ are the intensity values for R, G, and B channels, respectively.

After the conversion to gray scale, the previous frame I_{t-1} is subtracted from the current frame I_t to create an absolute valued difference $DI_{t,t-1}$, i.e.,

$$DI_{t,t-1} = |I_t - I_{t-1}| \quad (2)$$

Once the difference image is obtained, thresholding is applied to differentiate the moving pixels from the non-moving pixels. This process generates a binary difference image $BDI_{t,t-1}$ using a threshold value $\tau(DI_{t,t-1})$, the threshold value $\tau(DI_{t,t-1})$ is statistically obtained as

$$\begin{aligned} \tau(DI_{t,t-1}) &= \mu(DI_{t,t-1}) + \sigma(DI_{t,t-1}) \\ \mu(DI_{t,t-1}) &= \frac{\sum_{y=1}^H \sum_{x=1}^W DI_{t,t-1}(y,x)}{HW} \\ \sigma(DI_{t,t-1}) &= \sqrt{\frac{\sum_{y=1}^H \sum_{x=1}^W (DI_{t,t-1}(y,x) - \mu(DI_{t,t-1}))^2}{HW - 1}} \end{aligned} \quad (3)$$

Where $\mu(DI_{t,t-1})$ and $\sigma(DI_{t,t-1})$ are the mean and the standard deviation of the difference pixels, respectively, and (y, x) is the spatial coordinate. The definition of standard deviation given in (3) requires mathematical operations of square root and square. These operations are expensive in computation and require more energy consumption. Thus, we modify it as

$$\sigma(DI_{t,t-1}) = \frac{\sum_{y=1}^H \sum_{x=1}^W |DI_{t,t-1}(y,x) - \mu(DI_{t,t-1})|}{HW-1} \quad (4)$$

Using the threshold $\tau(DI_{t,t-1})$ the binary difference image $B DI_{t,t-1}$ is created as

$$B DI_{t,t-1}(y,x) = \begin{cases} 1, & \text{if } DI_{t,t-1}(y,x) \geq \tau(DI_{t,t-1}) \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

Fig.1 shows the binarization process of the frame difference using two consecutive RGB frames. Fig. 3(a) and (b) shows the current and previous frames, respectively. The binary difference image according to (2) is shown in Fig. 3(c). The thresholding process according to (5) generates a binary image as shown in Fig. 3(d) for further processing. Generating a binary difference image using the interframe difference technique only requires a single frame memory to store the previous frame.

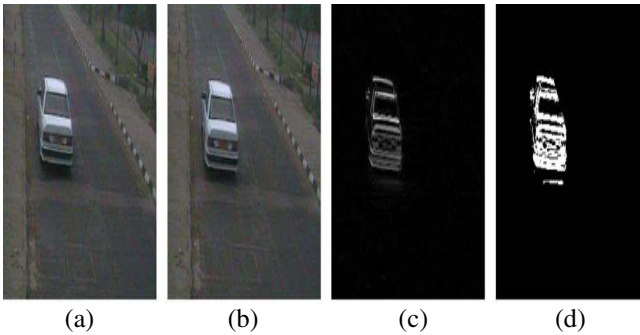


Fig. 1. Illustration of generating difference and binary difference images using the interframe difference technique. (a) Current RGB image. (b) Previous RGB image. (c) Difference image using (2). (d) Binary difference image using (5).

2.2 Background Subtraction Technique

Background subtraction detection is the essential function for all surveillance system based on computer vision. Accurate background subtraction is a key element to successful object tracking. Background subtraction mainly used to determine region of foreground object. The basic method is by pixel-by-pixel absolute differentiation of consecutive video frame with background image. However, in practical usage for live traffic monitoring system, background tends to be dynamic. Background can be affected by variation of lightning conditions, camera vibration and surroundings conditions. Thus, several methods were proposed for background modelling technique.

Most popular method of background subtraction is modelling through statistic

method. One of the approaches is Gaussian Mixture Model probability distribution for each pixel. In this method mean and variance is updated with the pixel values from the new frames of video sequence. After few frames, the model has acquired enough information and decision is made for each pixel whether it's background or foreground. Background and foreground decision is based on parametric equation. Interframe difference-based binary image generation produces an abstract representation of moving objects. A better approach is to perform background subtraction which identifies moving objects from the portion of a video frame that significantly differs from a background model. It involves comparing an observed image with an estimate of the background image to decide if it contains any moving objects. The pixels of the image plane where there are significant differences between the observed and estimated images indicate the location of the objects of interest. Let us assume that the background image B_t of a specific scene is constructed. The current greyscale image I_t at time t is subtracted from the background image B_t as

$$BS_t = |B_t - I_t| \quad (6)$$

To find the absolute-valued difference image between I_t and B_t . The difference image is binarized using the same approach in (5), i.e., $BBS_t(y, x) = 1$, if

$$BBS_t(y, x) = \begin{cases} 1, & \text{if } BS_t(y, x) \geq \tau(BS_t) \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

Where (BS_t) is a statistical threshold that is calculated according to (3), and BBS_t is a binary difference image.

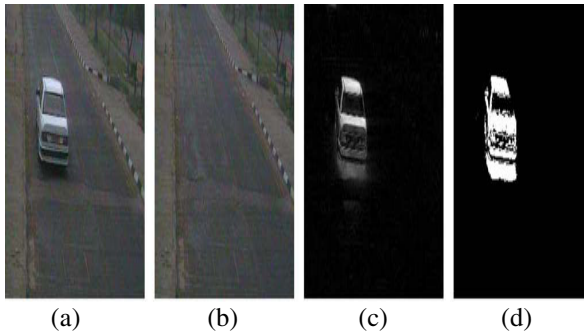


Fig. 2. Illustration of generating difference and binary difference images using the background subtraction technique. (a) Current RGB image. (b) Background RGB image. (c) Difference from the background image using (6). (d) Binary difference image using (7).

Fig. 2 shows the binarization process of the background subtraction technique. Fig. 2(a) and (b) shows the current and the background images, respectively. The difference image according to (6) is shown in Fig. 2(c). The thresholding process according to (7) generates a binary image that is shown in Fig. 2(d) for further processing.

The background image is updated according to BBS_t using BS_t and I_t as

$$B_{t+1}(y, x) = \alpha B_t(y, x) + (1 - \alpha) I_t(y, x), \text{ if } BBS_t(y, x) = 0$$

$$B_t(y, x) \text{ otherwise}$$
(8)

where B_{t+1} is the updated background image, and $\alpha \in [0, 1]$ is an update factor that is settled as $\alpha = 0.95$ in this paper.

The mean filter method for background generation is used because the background in the original color is needed and also it is an effective way for generating background, particularly, in low light variation and stationary camera posture, as it produces accurate results. The equation used in this study is described in the following formula

$$K_{xy} = \frac{\sum_{m=0}^{j-1} f_{xy}(t_{n-m})}{j}$$
(9)

Where n is frame number $f_{xy}(t_n)$ is the pixel value of (x,y) in n 'th frame $k_{xy}(t_n)$ is the pixel mean value of (x,y) in n 'th frame averaged over the previous j frames and j is the number of the frames used to calculate the average of the pixels value.

3 Speed Detection

The speed of the vehicle in each frame is calculated using the position of the vehicle in each frame, so the next step is to find out the blobs bounding box, and the centroid. The blob centroid is important to understand the distance of the vehicle moving in consecutive frames and therefore as the frame rate of captured moves is known, the calculation of the speed become possible. This information must be recorded consecutively into an array cell in the same size as the captured camera image because the distance moved by the centroid is needed which is a pixel with a specific coordinate in the image to find out the vehicle speed. To find out the distance moved by the pixel, suppose the pixel has the coordinate like

$$i = (a, b) \quad i-1 = (e, f)$$
(10)

Where the centroids location is showed in frame i and $i-1$ for one vehicles, with (a, b) coordinate and (e, f) coordinate. The distance difference for the vehicle is equal to

$$d1 = \sqrt{(a-e)^2 + (b-f)^2}$$
(11)

and if the image sequence is 25 frames per second, the time between two consecutive frames is equal to 0.04 s and finally the speed can be determined from the equation.

$$v = K \frac{\Delta_x}{\Delta_y}$$
(12)

Where K is the calibration coefficient.

Speed Violated Vehicle Detection Using Shrinking Algorithm

The speed estimation process is related with the tracking objects [5] in binary difference image BI_t where $BI_t \in \{BI_{t,t-1}, BBS_t$, the tracking and speed estimation using BI_t , consists of the following steps and configuration for the speed measurement of a moving vehicle shown in figure 3

- (1) Use the binary image BI_t and segment it into groups of moving objects using the aforementioned shrinking algorithm to create FR_t 's over region R_0 .
- (2) Track each FR_t in consecutive frames and find its spatial bounding box coordinates, i.e., upper left side coordinate of the spatial bounding box (y_t, x_t) at time instant t .
- (3) Trigger the timing t_i when the object passes the first imaginary line located at y_1 i.e., $y_{ii} \leq y_1$ and record its upper left side coordinate of the spatial bounding box, i.e., (y_{ii}, x_{ii}) .
- (4) Trigger the timing t_e when the object passes the second imaginary line located at y_2 , i.e., $y_{ie} \leq y_2$ and record its upper left side coordinate of spatial bounding box, i.e., (y_{ie}, x_{ie}) .
- (5) Estimate the speed of moving the vehicle by $v = \frac{\Delta y}{\Delta x} = \frac{y_{ie} - y_{ii}}{t_e - t_i}$
- (6) If the speed V is lower than the speed limit, then discard the object and go to step 1.
- (7) Extract the license plate using color information.
- (8) Transmit the extracted license plate image to the authorized remote station.
- (9) Go to step (1).

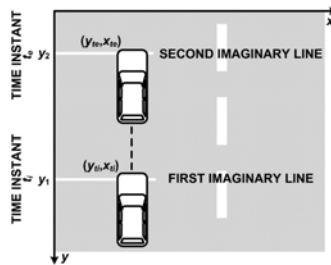


Fig. 3. Configuration for the speed measurement of a moving vehicle

4 Results

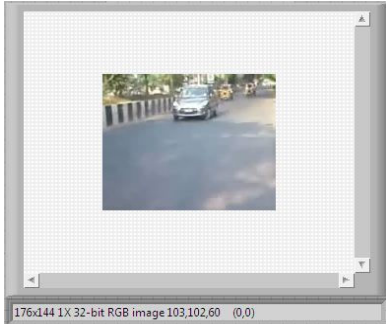


Fig. 4. RGB Image



Fig. 5. Background Image



Fig. 6. Grayscale image

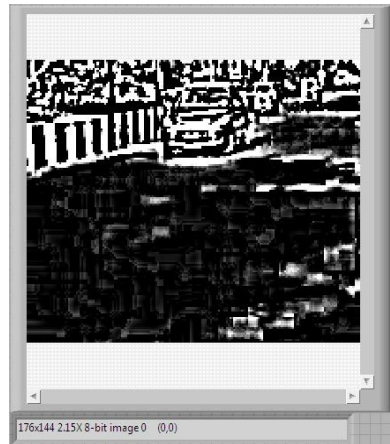


Fig. 7. Binary Image

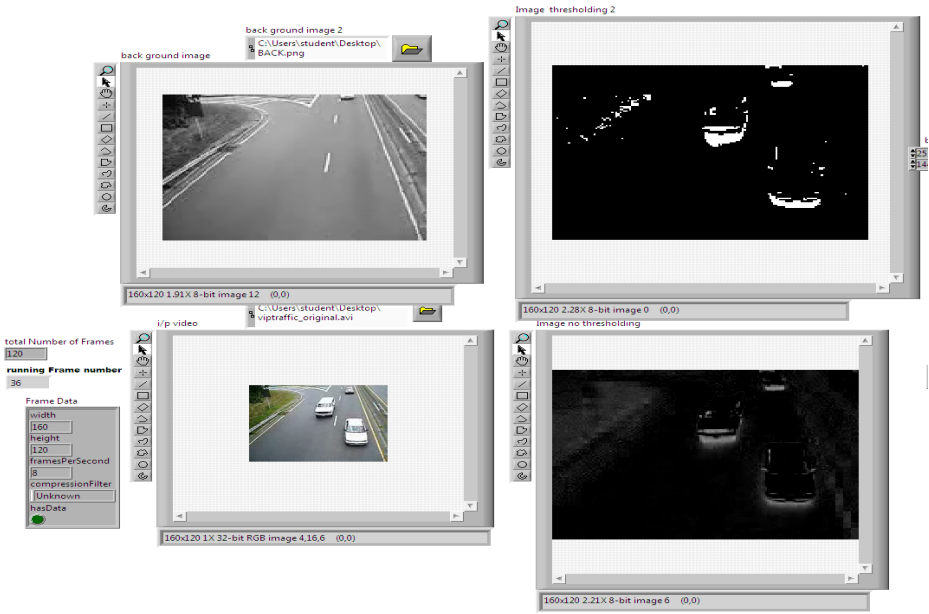


Fig. 8. LabVIEW front panel

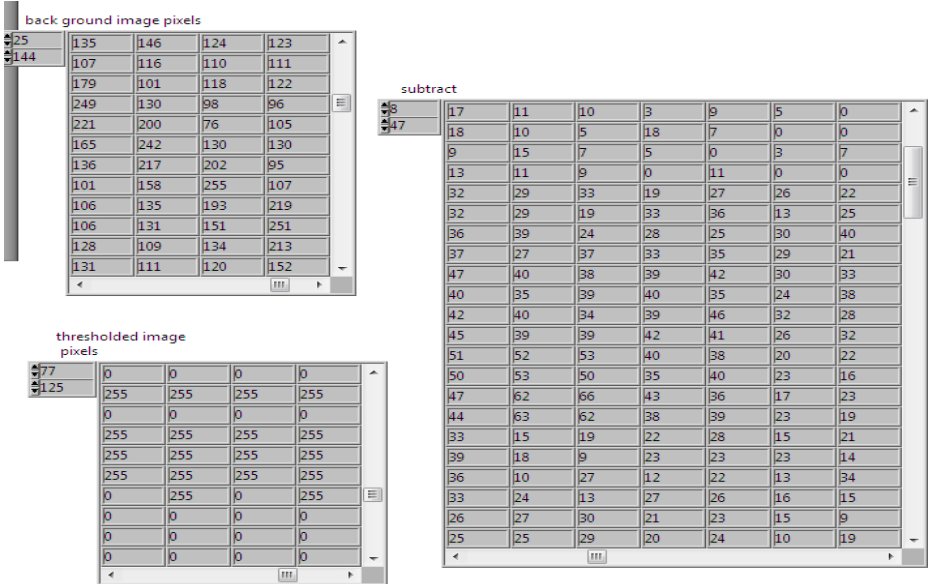


Fig. 9. Pixel values for image subtraction



Fig. 10. LabVIEW Front panel for vehicle speed detection

Table 1. Vehicle Speed detection using Shrinking algorithm

Vehicle number	True speed (km/h)	Interframe difference		Background subtraction	
		Estimated speed (km/h)	Error (km/h)	Estimated speed (km/h)	Error (km/h)
1	59.60	61.73	2.13	60.72	1.12
2	82.80	80.48	2.32	83.58	0.78
3	91.30	92.58	1.28	92.83	1.53
4	79.00	79.80	0.80	79.69	0.69
5	64.60	63.25	1.35	65.76	1.16
6	83.30	84.48	1.18	85.00	1.70
7	105.30	104.54	0.76	107.68	2.38
8	98.70	100.16	1.46	99.83	1.13
9	63.20	65.94	2.74	63.64	0.44
10	86.70	86.98	0.28	88.07	1.37
		Average Error	1.43	Average Error	1.23

5 Conclusions

In this paper, we have presented a speed camera system for speeding vehicle detection. The speeding vehicles are detected using the image processing techniques over the input image sequence captured from the fixed-position video camera. The speed estimation of vehicles is obtained in real-time using frame differencing technique for moving vehicle detection. In order to measure the speed of a vehicle, several models are proposed in this paper. Firstly, the background subtraction method

with a proper background updating model is used to detect the moving vehicle. The proposed algorithm has been developed using frame differencing technique and the formulation of velocity has been derived from its mathematical equation. The speeding vehicles are detected and tracked in consecutive sequences the image of the overall speeding vehicle is reported to the authorized remote station. The accuracy of the proposed system in the speed measurements is comparable with the true speed of the moving vehicles. Best results are obtained in shinking algorithm.

Acknowledgement. This research was supported by grant from Department of TIFAC CORE in Telematics in V R Siddhartha Engineering College, Vijayawada.

References

- [1] Ki, Y.-K., Baik, D.-K.: Model for accurate speed measurement using double-loop detectors. *IEEE Trans. Veh. Technol.* 55(4), 1094–1101 (2006)
- [2] Cheng, H., Shaw, B., Palen, J., Lin, B., Chen, B., Wang, Z.: Development and field test of a laser-based nonintrusive detection system for identification of vehicles on the highway. *IEEE Trans. Intell. Transp. Syst.*
- [3] Cucchiara, R., Piccardi, M., Mello, P.: Image analysis and rule-based reasoning for a traffic monitoring system. *IEEE Trans. Intell. Transp. Syst.* 1(2), 119–130 (2000)
- [4] Canals, R., Roussel, A., Famechon, J.-L., Treuillet, S.: A biprocessororiented vision-based target tracking system. *IEEE Trans. Ind. Electron.* 49(2), 500–506 (2002)
- [5] Cucchiara, R., Grana, C., Piccardi, M., Prati, A.: Detecting moving objects, ghosts, and shadows in video streams. *IEEE Trans. Pattern Anal. Mach. Intell.* 25(10), 1337–1342 (2003)
- [6] Schoepflin, T., Dailey, D.: Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation. *IEEE Trans. Intell. Transp. Syst.* 4(2), 90–98 (2003)
- [7] Gupte, S., Masoud, O., Martin, R., Papanikolopoulos, N.: Detection and classification of vehicles. *IEEE Trans. Intell. Transp. Syst.* 3(1), 37–47 (2002)

Feature Extraction from Lunar Images

V. Tamililakkiya and K.Vani

Information Science and Technology,
College of Engineering, Guindy,
Anna university-600025
ilakkiyamtech@gmail.com,
vani@annauniv.edu

Abstract. Automatic feature identification from orbital imagery would be of wide use in planetary science. For geo scientific applications, automatic shape-based feature detection offers a fast and non-subjective means of identifying geological structures within data. Most previously published examples of circular feature detection for geo scientific applications aimed to identify impact craters from optical or topographic data. Various techniques used include the texture analysis, template matching, and machine learning. In this paper, we propose a new method for the extraction of features from the planetary surface, based on the combination of several image processing techniques, including a shadow removal, watershed segmentation and the Circular Hough Transform (CHT). The original edge map of craters is detected by canny operator. In most literatures Hough transform is generally used for crater detection but we have added a shadow removal which includes a novel color image fusion method, based on the multi-scale Retinex (MSR) and discrete wavelet transform (DWT), is proposed. This proposed method is capable of detecting partially visible craters, and overlapping craters.

Keywords: feature extraction, shadow removal, multi-scale Retinex, discrete wavelet transform, image segmentation, circular hough transform.

1 Introduction

Craters are commonly found on the surface of planets, satellites, asteroids and other solar system bodies. The number and size of the crater is used to identify age of the surface which has craters. The relative ages between the surfaces of different bodies can also be identified. Generally, the crater is a bowl shaped depression created by collision or volcanic activities. Craters may have sharper and regular rims when they are younger and when the age of the crater increases regular rims leads to vague rims. Therefore, the craters are a fundamental tool to determinate the geological age and history of a surface.

Lunar images have features like rocks, craters, and ridges. Impact craters are objects that are caused by two celestial bodies impacting each other, such as a meteorite hitting a planet. Rocks are objects of small elliptical or circular shape, with no shadows present, and ridges appear like curves and straight lines in the images. Extraction of these features from lunar image by a human expert is a difficult task and

time consuming endeavour, because planetary images are blurry, quite noisy, present lack of contrast, the quality of the images is generally low (i.e., it depends on illumination, surface properties and atmospheric state), the features that are present in the images can be barely visible due to erosion and the shadows present in the images, and uneven illumination, and the represented objects are not well defined. Therefore, it is highly desirable procedure to detect the position, structure, and dimension of each feature is a trustworthy automatic procedure.

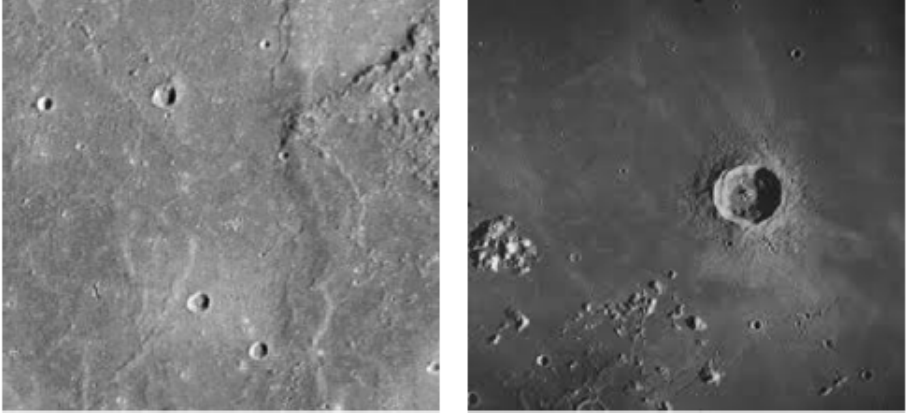


Fig. 1. (a)(b)-Lunar image with craters

Figure 1 shows the Lunar image with impact craters. These features also exhibit different structures and variable sizes.

The wide range of crater dimensions (from a few meters to thousands of kilometers) with distinct conservation conditions (from very fresh and well contrasted to very old with vague rims and filled or covered by other geological materials) occurring in quite diverse geo morphological settings has also made difficult the choice of adequate and unique parameters on the different automated approaches.

The proposed approach, aimed at extracting planetary features. Related works are discussed in Section 2. Shadow removal is discussed in Section 3 Pre-processing is done in Section 4, Followed by canny edge detection and crater detector in Section 5 and 6. The presentation and analysis of the results is included in Section 7. Finally, conclusions and ideas for future improvements are presented in Section 8.

2 Related Work

Some have employed texture analysis for crater detection[1]. They calculate variance in each tile and use shadows (high local variance) to detect crater or rocks. Still other algorithms are used for identification of craters. However, the results presented were far from being good enough to convince the deep space exploration research group to adopt many of these approaches.

Multiple boundary based approaches and merged to obtain the results [2]. In the first approach they have used images that are classified considering illumination.

When shady and illuminated pattern is recognized, they fit a circle to the surrounding edges. Although abrupt brightness changes may reveal a lot about the surface under consideration, the presence of sensor data with correct illumination is often an unrealistic assumption to make. The second approach they have used tries to find edge pixels of interest using a vectorized feature extractor. Then a roundness measure is checked for identification of circles.

Wetzler et al.[3] have used various supervised learning algorithms, including ensemble methods (bagging and AdaBoost with feed forward neural networks as base learners), support vector machines (SVM), and continuously scalable template models (CSTM) to derive crater detectors from ground truthed images. They have noted that the SVM solution to the problem performs superior on crater detection and localization compared to boundary-based rotation symmetry in the input image. This allows partially circular features to be picked up, and provides a natural way of fusing the results from multiple data sources. Subsequent to the detection of potential crater candidates in multiple data sets further analysis can be performed to remove circular noncrater-like features from consideration.

Two approaches proposed by Sawabe et al.[10] uses Hilditch's thinning algorithm and fuzzy Hough transform in addition to previously discussed algorithms respectively. These above approaches up to now suffer from elliptic shape of impact craters. Depending on geological deformations on the surface, there is a high possibility that craters form degraded craters rather than circles on the surface.

In this paper circular hough transform is used for crater detection. Before crater detection, removal of shadows present in the image is done. Which uses the MSR to enhance the color images can achieve adequate enhancement of shadowed detail and fails to rendition. The good color rendition and more details in shadows derived from the source image and the enhanced image can be combined into a fused image by DWT. Using this method, the fused image not only can preserve more details in shadows but also is more adapt the human visual perception. With these method more craters has been detected then crater detection without shadow removal.

3 Shadow Removal

Because of the underexposure or illumination, there are some shadows in the image it results in invisibility of the details or incapable distinguish by human vision. To stand out more details the single-scale Retinex (SSR) and multi-scale Retinex(MSR) algorithm based on the Retinex theory can achieve the dynamic compression. So, for eliminating shadows the Retinex algorithm has been widely applied to the remote sensing, medical image, removing haze from the image. After the source image with more dark areas is enhanced using SSR or MSR, the resulted image color is weaker and have the more details in shadow. Furthermore, if the color of the image is restored, the partly details in shadow are invisible too, accordingly the information is lost. So, the fusion method is proposed, in order to solve this problem.

A. Color Image Fusion

The fusion method proposed by this paper is performed based on the only one single image with many shadows in which the details is invisible. First, the original image is enhanced by MSR using auto-stretching method, in order to more details in shadow is

visible. Then, using DWT the source image and the enhanced image are fused. The color image fusion method can be performed in the following steps:

- 1) Color space of the original image and the enhanced image using MSR into the IHS color space transformation is made.
- 2) Using DWT decompose the two intensity components respectively.
- 3) Fusion rule: for detail coefficients select the maximum absolute value and average between the approximate coefficients.
- 4) By inverse DWT the fused coefficients are reconstructed into the new intensity component V' .
- 5) Consider the V' as the intensity component of the original image, and transform the new intensity, the saturation, and the hue, components back into RGB color space.

4 Preprocessing

Removal of shadows present in the image is done before crater detection, which uses the MSR to enhance the color images. It can achieve adequate enhancement of shadowed detail and fails to rendition. The DWT combines the good color rendition and more details in shadows derived from the source image and the enhanced image into a fused image by DWT. Using this method, the fused image not only can preserve more details in shadows but also is more adaptable to the human visual perception. The resulting image of shadow removal is shown in Figure2.

The 5*5 median filter and Gaussian filter is used for noise reduction and the original image is blurred as the small size of craters can't be extracted synchronously. The median filter is a nonlinear digital filtering technique, used to remove noise. Such noise reduction is a pre-processing step to improve the results of later processing. Median filtering is very widely used in digital image processing as it preserves edges while removing noise.

Gaussian filter is a filter for which impulse response is a Gaussian function. Gaussian filters are designed to give no overshoot to a step function input as they minimizing the rise and fall time. In two dimensions, it is said to be the product of two such Gaussians, one per direction:

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\pi\sigma^2}}$$

where x is the distance from the origin in the horizontal axis, y is the distance from the origin in the vertical axis, and σ is the standard deviation of the Gaussian distribution. When applied in two dimensions, above formula produces a surface in which contours are concentric circles with a Gaussian distribution from the center point.

Values from this distribution are used to build a convolution matrix and is applied to the original image. Each pixel's new value is set to a weighted average of that pixel's neighborhood. The original pixel's value receives the heaviest weight and the neighboring pixels receive smaller weights as their distance to the original pixel increases. Figure 3 shows the result of median and Gaussian filtering.

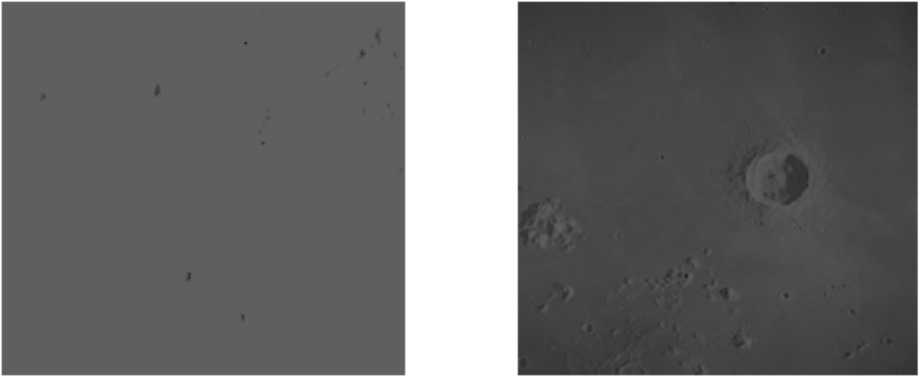


Fig. 2. (a),(b)-Shadow removal image

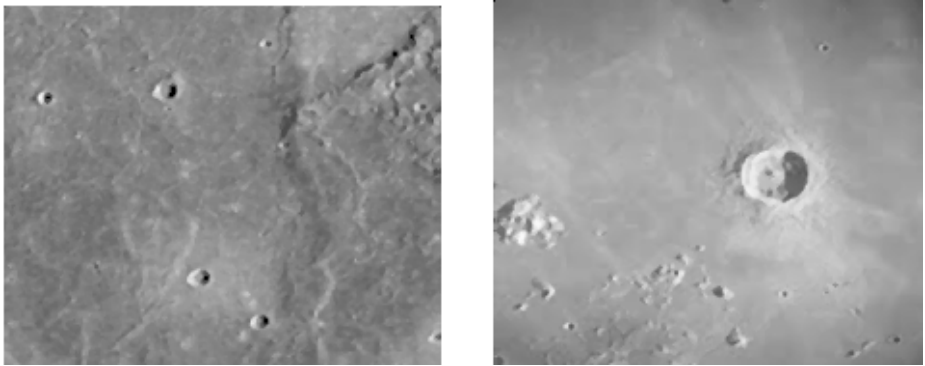


Fig. 3. (a),(b) -Filtered image using median and gaussian filter

Steps involved in crater detection:

- Shadow are removed based on the multi-scale Retinex (MSR) and discrete wavelet transform (DWT)
- The noise is smoothed by applying a Gaussian filtering and a median filtering operation.
- Compute image gradient I_g , by using the canny edge detector.
- By applying a non-maximum suppression algorithm followed by an Hysteresis threshold to I_g , a binary gradient image I_b , shows the contours of the objects is obtained.
- A circular Hough transform (CHT) is used to identify the seed points to detect these structures from I_b .
- For every pair of points that are detected as edge points in I_b and exhibit opposite gradient directions, an accumulator, corresponding to the median point between them, is incremented of a unit value.
- The maxima of the accumulator are taken as centers of the circle

5 Crater Edge Detection Using Canny

Canny operator was used to discover the optimal edge detection algorithm. Once this process is complete we have a binary image $I(x, y)$ where each pixel is marked as either an edge pixel or a non-edge pixel. From complementary output of the edge tracing step, the binary edge map obtained in this way can also be treated as a set of edge curves, which after further processing can be represented as polygons in the image domain. Then, to detect the edges, the image gradient is computed by using the Canny edge detector. Result of a canny edge detector is shown in figure 4.

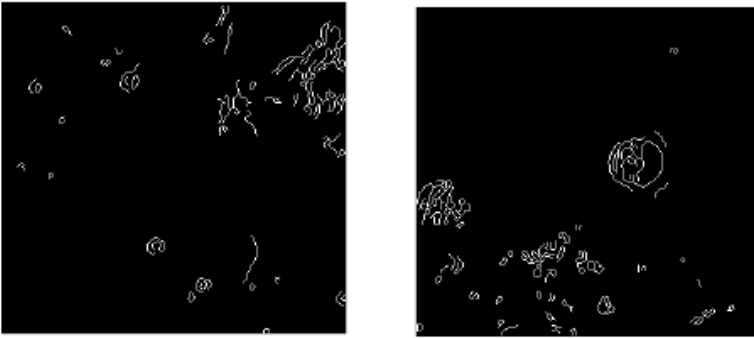


Fig. 4. (a)(b) Detection of edges using canny operator

As an intermediate result of the operation an intensity gradient I_g is generated, which is a gray-scale image; then, by applying a non-maximum suppression algorithm followed by a hysteresis thresholding to I_g , a binary gradient image I_b which shows the contours of the objects represented in the original image is obtained.

6 Crater Detection

Rocks, craters, and ridges are the features that are to be extracted. Crater in an image has an elliptical rim or circular rim and a bright to dark shading pattern. Inside of it, the image intensity profile along with lighting direction should be a monotonously decreasing function. Rocks are objects of small elliptical or circular shape, with no shadows, and the ridges appear like curves and straight lines in the images. Crater may have shadows present in an image, but the crater differs from rocks in this feature. Extraction of these features is a difficult task, because planetary images are blurry, quite noisy, present lack of contrast and uneven illumination, and the represented objects are not well defined. In order to address such problem a region-based approach, based on segmentation, has been chosen.

Shapes that are not clear may not be easily expressed using a small set of parameters. We must explicitly list the points on the shape for these cases. Make a table that contains all the edge pixels for our target shape. Position relative to some of the reference point for the shape can be stored for each pixels.

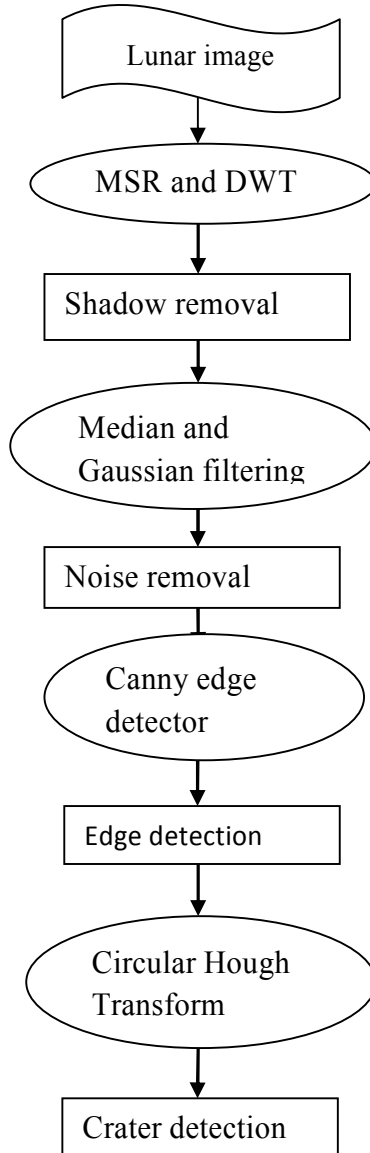


Fig. 5. Architecture diagram

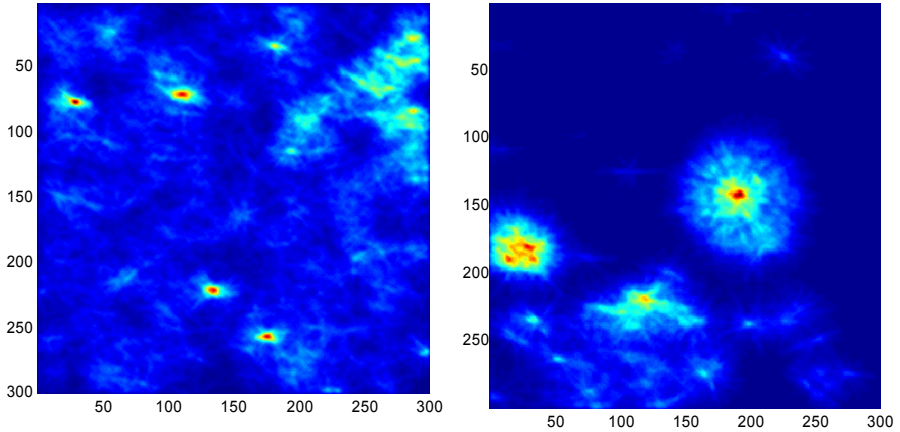


Fig. 6. (a)(b)-accumulation array of hough transform

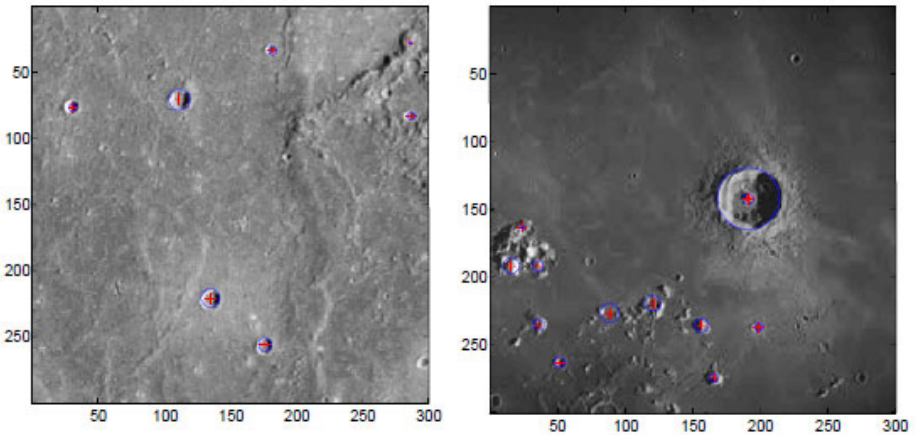


Fig. 7. (a)(b) detection of craters using circular Hough transform

Circular Hough Transform and can be expressed as follows:

1. Find all the desired feature points in the image
2. For each of the feature point
3. For each of the pixel i on the target's boundary
4. Get the relative position of reference point from i
5. Add the offset to the position of i
6. Increment this position in the accumulator
7. Find the local maxima in the accumulator
8. If it is desired, map each maxima in the accumulator back to image space using the target boundary table.

In order to identify the seed points that detect the structures from binary gradient image, I_b , the Circular Hough Transform is used. The median point between every pair of points

that are detected as edge points in I_b and exhibit opposite gradient directions, an accumulator, is incremented by a unit value. The maxima of the accumulator are taken as the centers of the craters or circles. Over all architecture diagram is shown in figure 5. Accumulation array of circular Hough transform is shown in figure 6.

Estimation of the two semi-axes and the direction angle of the crater from all of the pairs of points that contributed to the accumulator in the considered center is achieved. A seed point for segmentation is the center of the crater which has been generated truly corresponds to a contour in the gradient image.

Starting from all the detected seed points, the 3D accumulator are used to identify craters by applying the watershed algorithm to gradient image, I_g . By using the parameters describing the crater centered in each detected maximum are computed.

Target range of crater diameter for recognition is an important value in this recognition algorithm, because the vote for big craters gets priority if the range of recognition is not decided. Small craters can be recognized by defining the target range. Therefore, the target range of crater diameter should be limited for the effective crater recognition. Experimental results of crater detection show that the CHT can find craters with incomplete rims, nested craters, and craters overlaid on the rim of another crater.

7 Results and Discussions

We tested the proposed system on a set of 20 images, with different characteristics in terms of size, noise, and content. The results we obtained were compared against the craters extracted by a human operator, who was asked to pick up the craters both by looking at the grey level image and the accumulator. As a result most of the craters were extracted from grey level images and accumulator, which is shown in figure 7.

To give an idea of the capability of the algorithm to extract craters even in very complex images where several craters are present along with many spurious edges and partially visible and overlapped. With shadow removal method using MSR and DWT this partially visible and overlapped craters are identified. Since radius range is specified it fails to detect very large size of craters that are not present within the range.

8 Conclusion

This paper describes a circular Hough transform for impact crater detection which accounts for changes in illumination, visual appearance, and size. The results obtained outperform most published approaches. There are many shadows in which details are invisible. After the source image is enhanced by MSR, the enhanced image and the source image are fused into a fused image. The fused image has more information and its color is more adapt to the human visual system. The circular Hough transform with shadow removal is able to automatically select a small set of features characterizing the presence/absence of craters in an image. However, by increasing radius range large size craters can also be identified. The further work should improve the choice method of seed and use some intelligent methods to decide the size of crater candidate area, and plan to evaluate the performance of this algorithm on other terrains of Mars and also on other planetary surfaces.

References

1. Barata, T., Alves, E.I., Saraiva, J., Pina, P.: Automatic Recognition of Impact Craters on the Surface of Mars. In: Campilho, A.C., Kamel, M.S. (eds.) ICIAR 2004. LNCS, vol. 3212, pp. 489–496. Springer, Heidelberg (2004)
2. Sugiyama, T., Abe, K.: Edge feature analysis by a vectorized feature extractor and in multiple edge. IEIC, D-2, J80-D-2 6, 1379–1389 (1997)
3. Wetzler, P.G., Honda, R., Enke, B., Merline, W.J., Chapman, C.R., Burl, M.C.: Learning to Detect Small Impact Craters. In: Proc. of the Seventh IEEE Workshops on Application of Computer Vision, vol. 01 (2005)
4. Kim, J.R., Muller, J.P., Gasselt, S.V., Morley, J.G., Neukum, G.: Automated Crater Detection, A New Tool for Mars Cartography and Chronology. *Photogrammetric Engineering and Remote Sensing* 71(10), 1205–1217 (2005)
5. Honda, R., Azuma, R.: Crater Extraction and Classification System for Lunar Images, Technical report at Department of Mathematics, Kochi University (2000)
6. Leroy, B., Medioni, G., Johnson, E., Matthies, L.: Crater detection for autonomous landing on asteroids. *Image and Vision Computing* 19, 787–792 (2001)
7. Bue, B.D., Stepinski, T.F.: Machine Detection of Martian Impact Craters From Digital Topography Data. *IEEE Trans. Geosci. and Remote Sens.* 45, 265–274 (2007)
8. Bandeira, L., Saraiva, J., Pina, P.: Impact Crater Recognition on Mars Based on a Probability Volume Created by Template Matching. In: TGRS-2007, pp. 1–8 (2007)
9. Earl, J., Chicarro, A., Koeberl, C., Marchetti, P.G.: Automatic recognition of crater-like structures in terrestrial and planetary images. *Lunar and Planetary Science XXVI* (2005)
10. Sawabe, Y., Matsunaga, T., Rokugawa, S.: Automated detection and classification of lunar craters using multiple approaches. *Adv. Space Res.* 37(1), 21–27 (2006)
11. Zhang, X.: A Novel Color Image Fusion Method Based on the Multi-Scale Retinex and DWT. In: ICIME, pp. 395–398 (2009)

A Hybrid Approach to Estimate True Density Function for Gene Expression Data

Ganesh Kumar Pugalendhi¹, Mahibha David², and Aruldoss Albert Victoire³

¹ Assistant Professor , Department of IT, Anna University of Technology, Coimbatore, India
ganesh23508@gmail.com

² Assistant Professor, Department of CSE, ² SCAD College of Engg. and Tech., India
mahibha_it2007@yahoo.co.in

³ Assistant Professor, Department of EEE, Anna University of Technology, Coimbatore, India
t.aruldoss@gmail.com

Abstract. Accurate classification of diseases from microarray gene expression profile is a challenging task because of its high dimensional low sample data. Most of the gene selection methods discretize the continuous-valued gene expression data for estimating the marginal and joint probabilities that results in inherent error during discretization and reduces the classification accuracy. To overcome this difficulty, a hybrid fuzzy-rough set approach is proposed that generates a fuzzy equivalence class and constructs a fuzzy equivalence partition matrix to estimate the true density function for the continuous-valued gene expression data without discretization. The performance of the proposed approach is evaluated using six gene expression data. f -Information measure is used for gene selection and back propagation network is used for classification. Simulation results show that the proposed method estimate the true density function correctly without discretizing the continuous gene expression values. Further the proposed approach performs the integration required to compute f -Information measure easily and results in highly informative genes that produces good classification accuracy.

Keywords: Gene Expression profiles, Fuzzy-Rough Set, f -Information, Back Propagation Network.

1 Introduction

DNA microarrays [1] are an important technology for studying gene expression. With a single hybridization, the level of thousands of genes, or even entire genome, can be estimated from a sample of cells. Generally the microarray data are images, which have to be transformed into gene expression matrices in which rows represent genes, columns represent various samples such as tissues or experimental conditions, and numbers in each cell characterize the expression level of the particular gene in the particular sample. Recently gene expression profiles are more preferred form of disease diagnostic system than morphology [2]. The number of genes in a gene expression profile (usually in the range of 2,000-30,000) is much larger than the number of samples (usually in the range of 40-200). This high dimensional low sample

data poses a lot of difficulties when it is analyzed by manual. Hence the need to automatically analyze the microarray data has a significant impact on diagnosing disease from gene expression profiles.

A diagnostic system designed using the large set of gene expression features will have higher computational cost, slower learning process and poor classification accuracy due to the phenomenon known as curse of dimensionality. Recently researches [3] have shown that a small number of genes are sufficient for accurate diagnosis of most of the diseases. More importantly, by using small subset of genes, we can not only get a better diagnostic accuracy, but also get an opportunity to further analyze the nature of the disease and the genetic mechanisms responsible for it. Thus dimensionality reduction plays a major role in selecting informative genes from microarray gene expression data.

There are two different approaches to achieve dimensionality reduction [10]: Feature extraction and feature selection. Feature extraction linearly or nonlinearly transforms the original set of genes into reduced one. Feature selection selects a subset of genes from the initial set of available genes which leads to savings in measurement cost and the selected genes retain their original physical interpretation. On the other hand transformed genes generated by feature extraction may not have a clear physical meaning and are very hard to interpret. Hence feature selection is most preferred for a gene expression profile based diagnostic system.

Feature selection [13] methods can be classified into three categories depending on how they combine the feature selection search with the construction of the classification model. If feature selection is performed independently of the learning algorithm, the technique is said to follow a filter approach. In contrast wrapper method incorporates the learning algorithm in the feature selection process. The third class of feature selection techniques termed as embedded technique search for an optimal subset of feature is built into the classifier construction and can be seen as a search in the combined space of features subsets and hypotheses. Because of the high computational efficiency, filter methods are very popular to high dimensional data and seems to be an appropriate method in selecting informative genes from high dimensional low sample gene expression profile.

So far, lots of filter based gene selection methods have been proposed to identify informative genes from gene expression profiles. A common practice in filter type method is to employ a criterion function to evaluate the relevance or discriminant power of the genes with regard to target diseases. Chi-squared, entropy [4], [6], mutual information [5], f -information [7] are some filter based feature selection techniques that arrange the genes according to their predictive power. All the above mentioned information measures divide the continuous expression values of a gene into several discrete partitions. The marginal probabilities and their joint probabilities are then calculated to measure both gene-class relevance of a gene with respect to class labels and gene-gene redundancy between two genes.

But the inherent error that occurs in the discretization process is of major concern in the computation of information measures of continuous gene expression values. Histograms [8] are used to estimate the true density functions, and the computational complexity of performing integration can be overcome in an efficient manner. However the histogram based approach is only applicable to relatively low dimensional data as the sparse data distribution degrade the reliability of histograms. To overcome this difficulty a new hybrid approach is proposed that works directly

with the continuous real valued gene expression datasets for estimating the true density function.

The proposed method generates a fuzzy equivalence partition matrix (FEPM) by combining the concept of fuzzy and rough set theory. The generated FEPM is then applied to f -information measure to calculate the relevance and redundancy between genes. The performance of the proposed approach is tested using three gene expression data sets viz., Colon cancer [14], Lymphoma [16], Rheumatoid Arthritis versus Osteoarthritis (RAOA) [15], Rheumatoid Arthritis versus health controls (RAHC) [17], Leukemia[1] and Prostate cancer[18]. Artificial Neural Network [12] trained by back propagation algorithm is used to evaluate the genes selected by the proposed approach.

The structure of the rest of this paper is described as follows. In Section 2, implementation issues of the proposed fuzzy-rough set approach for generation of FEPM are presented. Details of simulations conducted using six gene expression data sets and the results are reported in Section 3. Concluding remarks are given in Section 4.

2 Proposed Fuzzy- Rough Set Approach

A Rough set [9] is an approximation of a vague concept by a pair of precise concepts, called lower and upper approximations. A fuzzy-rough set is a generalization of a rough set, derived from the approximation of a fuzzy set in a crisp approximation space. Let $\langle U, A \rangle$ represent the fuzzy approximation spaces and X be a fuzzy subset of U . The fuzzy P-lower and P-upper approximation can then be defined as,

$$\mu_{\underline{P}X}(F_i) = \inf_x \{ \max \{ (1 - \mu_{F_i}(x)), \mu_x(x) \} \}, \forall i \quad (1)$$

$$\mu_{\overline{P}X}(F_i) = \sup_x \{ \min \{ (\mu_{F_i}(x)), \mu_x(x) \} \}, \forall i \quad (2)$$

where F_i represents a fuzzy equivalence class belonging to U/P , and $\mu_x(x)$ represents the membership of x in X . This fuzzy equivalence class generated for each gene can be used to construct fuzzy equivalence partition matrix.

2.1 Generation of FEPM Matrix

- 1) Read the gene expression dataset $G_{i \times j}$, where $i = 1, 2, \dots, m$, c ; m =number of genes, c =class label = $1, 2, \dots, n$; n = number of samples.
- 2) Calculate the mean value $\mu = \{\mu_1, \mu_2, \dots, \mu_m, c\}$ for each gene of all the samples and class label.
- 3) Generate two gene groups (High H, Low L) by comparing each gene value with respective mean values, so that

$$\begin{aligned} H &= \{\text{Genes having value greater than its mean}\} \\ L &= \{\text{Genes having value lower than its mean}\} \end{aligned}$$

- 4) Calculate the mean value of two gene groups for each gene,

$$\mu_L = \{\mu_{L_1}, \mu_{L_2}, \dots, \mu_{L_c}\} \quad \mu_H = \{\mu_{H_1}, \mu_{H_2}, \dots, \mu_{H_c}\}$$

- 5) The mean value calculated in step 3 is considered as the medium mean value,

$$\mu_M = \mu_{M_1}, \mu_{M_2}, \dots, \mu_{M_c}$$

- 6) Calculate the standard deviation for each mean values $\{\mu_L, \mu_M, \mu_C\}$

$$\begin{aligned}\sigma_L &= \{\sigma_{L_1}, \sigma_{L_2}, \dots, \sigma_{L_c}\} \\ \sigma_H &= \{\sigma_{H_1}, \sigma_{H_2}, \dots, \sigma_{H_c}\} \\ \sigma_M &= \{\sigma_{M_1}, \sigma_{M_2}, \dots, \sigma_{M_c}\}\end{aligned}$$

- 7) Calculate the membership value (π_L, π_H, π_M) on fuzzy approximation spaces for each gene $G_{i \times j}$,

$$\pi_L(G_{i \times j}, \mu_{L_i}, \sigma_{L_i}) = \begin{cases} 2(1 - \|G_{i \times j} - \mu_{L_i}\|)^2, & \frac{\sigma_{L_i}}{2} \leq \|G_{i \times j} - \mu_{L_i}\| \leq \sigma_{L_i} \\ 1 - 2(\|G_{i \times j} - \mu_{L_i}\|)^2, & 0 \leq \|G_{i \times j} - \mu_{L_i}\| \leq \frac{\sigma_{L_i}}{2} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

- 8) Calculate the positional values $(P_L^{G_{i \times j}}, P_M^{G_{i \times j}}, P_H^{G_{i \times j}})$ for each gene.

$$P_L^{G_{i \times j}} = \frac{\pi_{L_{i \times j}}}{\pi_{L_{i \times j}} + \pi_{M_{i \times j}} + \pi_{H_{i \times j}}} \quad (4)$$

- 9) Form FEP matrix $FP_{i=}$ $\begin{bmatrix} P_L^{G_{i \times j}} \\ P_M^{G_{i \times j}} \\ P_H^{G_{i \times j}} \end{bmatrix}$ for each gene.

2.2 V-Information Measure on the Fuzzy Approximation Spaces

- 1) Initialize $I_{inf} = 0, I_{rel} = 0, I_{rem} = 0, I_{red} = 0$.
- 2) Read the FEP matrix FP_i , where $i = 1, 2, \dots, m, c$.
- 3) Calculate relevance between each gene and class label using V- information

$$\begin{aligned}V_{rel}(G_{i \times j}, G_c) = & \left| \frac{1}{n} \sum_{j=1}^n (P_L^{G_{i \times j}} \cap P_L^{G_{c \times j}}) - \frac{1}{n^2} \sum_{j=1}^n P_L^{G_{i \times j}} \sum_{j=1}^n P_L^{G_{c \times j}} \right| + \\ & \left| \frac{1}{n} \sum_{j=1}^n (P_H^{G_{i \times j}} \cap P_H^{G_{c \times j}}) - \frac{1}{n^2} \sum_{j=1}^n P_H^{G_{i \times j}} \sum_{j=1}^n P_H^{G_{c \times j}} \right| + \\ & \left| \frac{1}{n} \sum_{j=1}^n (P_M^{G_{i \times j}} \cap P_M^{G_{c \times j}}) - \frac{1}{n^2} \sum_{j=1}^n P_M^{G_{i \times j}} \sum_{j=1}^n P_M^{G_{c \times j}} \right| \quad (5)\end{aligned}$$

- 4) Sort the genes in descending order of $V_{rel}(G_{i \times j}, G_c)$.
- 5) Store the top N_{rel} genes in I_{rel} and remaining genes $(m - N_{rel})$ to I_{rem} .
- 6) Calculate redundancy value $V_{red}(I_{rel_{x \times j}}, I_{rem_{x \times j}})$ between selected genes and each of the remaining genes,

$$\begin{aligned}
& V_{red}(I_{rel_{x \times j}}, I_{rem_{x \times j}}) = \\
& \left| \frac{1}{n} \sum_{j=1}^n \left(P_L^{I_{rel_{x \times j}}} \cap P_L^{I_{rem_{x \times j}}} \right) - \frac{1}{n^2} \sum_{j=1}^n P_L^{I_{rel_{x \times j}}} \sum_{j=1}^n P_L^{I_{rem_{x \times j}}} \right| + \\
& \left| \frac{1}{n} \sum_{j=1}^n \left(P_H^{I_{rel_{x \times j}}} \cap P_H^{I_{rem_{x \times j}}} \right) - \frac{1}{n^2} \sum_{j=1}^n P_H^{I_{rel_{x \times j}}} \sum_{j=1}^n P_H^{I_{rem_{x \times j}}} \right| + \\
& \left| \frac{1}{n} \sum_{j=1}^n \left(P_M^{I_{rel_{x \times j}}} \cap P_M^{I_{rem_{x \times j}}} \right) - \frac{1}{n^2} \sum_{j=1}^n P_M^{I_{rel_{x \times j}}} \sum_{j=1}^n P_M^{I_{rem_{x \times j}}} \right| \quad (6)
\end{aligned}$$

$$7) \quad I_{red} = \max(V_{rel}(G_{i \times j}, G_c) - \frac{1}{|rel|} \sum_{I_{rel}} V_{red}(I_{rel_{x \times j}}, I_{rem_{x \times j}})) \quad (7)$$

$$8) \quad I_{inf} = \{I_{rel}, I_{red}\} \quad (8)$$

3 Simulation Results

This section presents the details of simulation carried out using six gene expression datasets. Table 1 gives the information about the gene expression data set used in the experiment. All these data sets are publicly available and are two class gene expression profiles.

Table 1. Details of gene expression dataset

Dataset	Total Samples	No. of Genes	Class Labels	Class wise Samples
Colon cancer [20]	62	2000	Tumor	40
			Normal	22
Lymphoma [22]	45	4026	Germinal Centre B-Like (GCL)	23
			Activated B-Like (ACL)	22
			Rheumatoid Arthritis (RA)	22
RAOA [21]	31	18,432	Osteoarthritis (OA)	9
			Rheumatoid	18
RAHC[23]	33	4,701	Health controls	15
Leukemia[1]	45	7129	Germinal Centre B-Like(GCL)	22
			Activated B-Like(ACL)	23
Prostate[24]	33	12,627	Normal	24
			Tumor	9

The proposed approach is implemented in MATLAB 7.5 and executed in a PC with Intel Core 2 Duo processor with 2.60 GHz speed and 2 GB of RAM. The steps of implementing the proposed approach are illustrated for colon cancer data set. Expression values for some selected genes and samples of Colon cancer dataset are given in Table 2.

Table 2. Sample gene expression values of colon cancer dataset

Gene/Sample	S_1	S_2	S_{61}	S_{62}
H55933	8589.4	9164.2	...	6234.623	7472.01
R39465	5468.2	6719.5	...	4005.3	3653.934
...
H40891	83.522	44.472	...	32.687	49.862
R77780	28.701	16.773	...	23.26	39.63

At first, fuzzy equivalence class is computed for each gene using the steps 1 to 7 as discussed in section 2.1. The FEC computed for each gene is then used to generate fuzzy equivalence partition matrix (FEPM) by following the steps 8 and 9 of section 2.1. FEC and FEPM computed for the gene H55933 is given in table 3.

Table 3. Value of FEC and FEPM for gene H55933

Fuzzy Equivalence Class for gene H55933					
FEC	S_1	S_2	...	S_{61}	S_{62}
Low	0.0174	0	...	0.8223	0.2920
medium	0.8818	0.9554	...	0.2775	0.6739
High	0.9324	0.8741	...	0.9834	0.9943
Fuzzy Equivalence Partition Matrix for gene H55933					
FEPM	S_1	S_2	...	S_{61}	S_{62}
Low	0.0095	0	...	0.039	0.15
medium	0.5091	0.4778	...	0.472	0.51
High	0.4814	0.5228	...	0.133	0.33

As shown in table 3, FEC and FEPM are constructed for all the genes of colon cancer data set, then Gene-Class relevance is calculated using step 3 of section 2.2. Based on the relevance value, genes are ranked and top ranked genes are selected as highly relevant genes. After relevance calculation, Gene-Gene redundancy is calculated between the selected genes and the remaining genes of the data set using the step 6 of section 2.2. The redundancy value calculated is given in the Table 4 for the remaining genes except the genes selected based on relevance value.

Table 4. Gene-Gene Redundancy Value

Gene No	Gene ID	V_{red}
G_1	H55933	1.0271
G_2	R39465	1.0802
...
G_{1999}	H40891	1.0621
G_{2000}	R77780	1.0284

Then using the step 7 of section 2.2, I_{red} values are calculated and are ranked. Fig.1. shows the I_{red} values obtained for first 30 genes of colon cancer dataset. From this, it is evident that only a few genes are having significant information about the diseases and the remaining genes have very less amount of information.

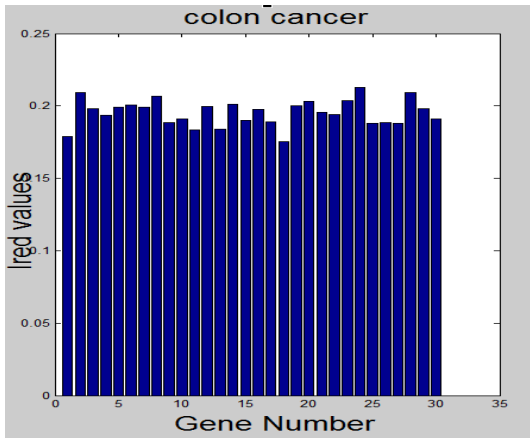


Fig. 1. I_{red} values of colon cancer data set

Finally, informative genes are selected based on the relevance and redundancy rank value. Thirty top ranked genes are selected as informative genes out of which 3 genes are top ranked relevant genes and the remaining 27 genes are top ranked redundant genes. Table 5 gives the details of some of the selected high informative genes based on the Gene-Class and Gene-Gene Framework.

Table 5. Informative Genes

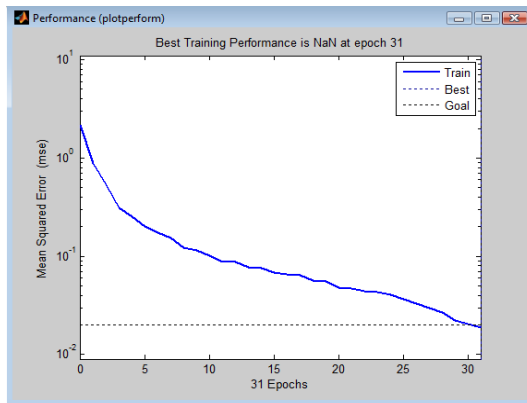
Gene No	Gene ID	Final Rank
G ₁₀₀₁	M16029	I ₁
G ₇₇₅	X13451	I ₂
G ₁₀₂₄	T67897	I ₃
G ₇₄₃	T59354	I ₄
G ₁₃₉₅	H10925	I ₅
G ₁₈₃	X51346	I ₆
G ₁₀₀₇	D11086	I ₇
G ₁₄₆₆	D00860	I ₈
G ₁₁₈₆	H40269	I ₉
G ₁₁₉	T72889	I ₁₀
G ₈₃₃	H29483	I ₁₁
G ₁₀₀₆	T53889	I ₁₂
G ₅₁₇	M59807	I ₁₃
G ₁₁₂₂	M22488	I ₁₄
G ₇₅₂	M32800	I ₁₅

The informative genes selected using the proposed v -information on fuzzy approximation space is evaluated by using them as input for the ANN based classifier model. The details of samples selected for training and testing the ANN model is given in Table 6.

Table 6. Details of sample partition

Gene expression dataset	Training	Total Training	Testing	Total Testing
Colon cancer	Normal:11 Tumor:20	31	Normal:11 Tumor:20	31
Lymphoma	GCL:12 ACL:11	23	GCL:11 ACL:11	22
RAOA	RA:11 OA:5	16	RA:11 OA:4	15
RAHC	RA:9 HC:8	17	RA:9 HC:7	16
Leukemia	ALL:24 AML:13	37	ALL:23 AML:12	35
Prostate cancer	Normal:5 Tumor:12	17	Normal:4 Tumor:12	16

At first, the ANN model is trained using the samples selected for training. Trial and error procedure was followed to identify the optimal number of hidden nodes. There are about 30 neurons in the input layer that corresponds to the informative genes and one neuron in the output layer that corresponds to class label. The number of hidden layers is directly related to the capabilities of the network. The network is trained with least means square error of 0.02. The mean square error achieved during training is 0.0189. With ten hidden nodes, the network took 1.9344 seconds to reach the error goal. The performance of the network during training is show in fig 2.

**Fig. 2.** Training performance of the network

After training, the generalization performance of the network is evaluated with the test data. During testing the mean square error achieved by the network is 0.3860 in 0.1248 seconds. The performance of the network during testing is presented in table 7 for all the data sets.

Table 7. Testing performance of the network

Gene Expression data	Nhid	Overall Accuracy	Normal Acc	Tumor Acc
Colon Cancer	10	90.32	81.81	95
Lymphoma	6	90.9	90.9	90.9
RAOA	1	93.33	90.9	100
Leukemia	12	97.14	95.65	100
RAHC	4	93.75	100	88.8
Prostate Cancer	1	87.5	50	100

From Table 7 it is found that the neural network can classify more number of samples especially the tumor samples for all the datasets. To show the power of the proposed Fuzzy-Rough set approach results are also taken with and without FEPM and it is reported in the Table 8.

Table 8. Performance of gene f -Information with and without FEPM

Dataset	With FEPM	Without FEPM
Colon cancer	90.32	87.09
RAOA data	93.33	86.6
Lymphoma	90.9	81.81
leukemia	97.14	74.28
RAHC	93.75	81.25
Prostate cancer	87.5	81.25

From this comparison it is found that the discretization error produced during gene selection without FEPM is reduced and the proposed method bring a remarkable improvement on the approximation of the true marginal and joint distribution of continuous valued data. The performance of the proposed approach is compared with other gene selection approaches and it is presented in Table 9.

Table 9. Performance comparison of proposed approach

Dataset	Gene selection method	Classification accuracy
Colon Cancer	SNR[10]	65
	EA[11]	75.8
	Proposed method	90.32
Lymphoma	SNR	76
	EA	74.47
	Proposed Method	90.9
RAOA	f -information[7]	86.6
	Proposed Method	93.33
RAHC	f -Information	81.25
	Proposed method	93.75
Leukemia	EA	72.4
	SNR	59
	Proposed method	97.14
Prostate Cancer	f -Information	81.25
	Proposed method	87.5

From this comparison, it is found that the proposed hybrid feature selection method has the ability to provide highly informative genes and achieves rich measure of classification performance than the other algorithms reported in the literature.

4 Conclusion

The bottleneck of analyzing gene expression data is the identification of informative genes and classification. A hybrid fuzzy rough set approach is proposed to construct a fuzzy Equivalence Partition matrix for approximating the true marginal and joint distribution of continuous gene expression values. f -Information based measure is used for identifying discriminative and non redundant genes from FEPM without discretization. Feed forward neural network trained by back propagation algorithm is developed for classification using the selected informative genes. Simulation results shows that the proposed approach is effective and accurate in microarray data analysis.

References

1. Simon, R.M., Korn, E.L., McShane, L.M., Radmacher, M.D., Wright, G.W., Zhao, Y.: Design and Analysis of DNA Microarray Investigations. Springer, Berlin (2003)
2. Golub, T.R., Slonim, D., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J., Coller, H., Loh, M., Downing, J., Caligiuri, M., Bloomfield, C., Lander, E.: Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 286, 531–537 (1999)
3. Hastie, T., Tibshirani, R., Eisen, M.B., Alizadeh, A., Levy, R., Chan, W.C., Bostein, W.C.D., Brown, P.O.: Gene Shaving as a method for identifying distinct set of genes with similar expression patterns. *Genome Biology* 1(2) (2000)
4. Ding, Peng, H.: Minimum Redundancy Feature selection from microarray gene expression data. *J. Bioinformatics Comput. Biol.* 3(2), 185–205 (2005)
5. Peng, H., Long, F., Ding, C.: Feature selection based on mutual information: Criteria of max-dependency, max-relevance. *IEEE Trans. Pattern Anal. Mach. Intell.* 27(8), 1226–1238 (2005)
6. Liu, X., Krishnan, A., Mondry, A.: An Entropy-based gene selection method for cancer classification using microarray data. *BMC Bioinformatics* 6(76), 1–14 (2005)
7. Maji, P.: f -Information Measures for Efficient Selection of Discriminative Genes From Microarray Data. *IEEE Trans. Biomed. Eng.* 56(4), 1063–1069 (2009)
8. Kwak, N., Choi, C.-H.: Input Feature Selection by Mutual Information Based on Parzen Window. *IEEE Trans. Pattern Anal. Mach. Intell.* 24(12), 1667–1671 (2002)
9. Maji, P., Pal, S.K.: Rough set based generalised fuzzy C-means algorithm and quantitative indices. *IEEE Trans. Syst., Man, Cybern. B, Cybern.* 37(6), 1529–1540 (2007)
10. Shi, C., Chen, L.: Feature dimension reduction for microarray data analysis using locally linear embedding. In: APBC, pp. 211–217 (2005)
11. Umpai, T.J., Aitken, S.: Feature selection and classification for microarray data analysis: Evolutionary methods for identifying predictive genes. *Bioinformatics* 6, 168–174 (2005)
12. Devaraj, D., Preetha Roseylyn, J., Umar Rani, R.: Artificial Neural model for voltage security based contingency ranking. *Int. J. on Applied Soft Computing* 7(3), 722–727 (2007)

13. Saeys, Y., Inza, I., Larranaga, P.: A review of feature selection techniques in bioinformatics. *Bioinformatics* 23(19), 2507–2517 (2007)
14. Alon, U., Barkai, N., Notterman, D.A., Gish, K., Ybarra, S., Mack, D., Levine, A.J.: Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc. Nat. Acad. Sci. U.S.A* 96(12), 6745–6750 (1999)
15. van der Pouw Kraan, T.C.T.M., van Gaalen, F.A., Kasperkovitz, P.V., Verbeet, N.L., Smeets, T.J.M., Kraan, M.C., Fero, M., Tak, P.-P., Huizinga, T.W.J., Pieterman, E., Breedveld, F.C., Breedveld, A.A., Alizadeh, A.A., Verweij, C.L.: Rheumatoid arthritis is a heterogenous disease: Evidence for differences in the activation of the STAT-1 pathway between rheumatoid tissues. *Arthritis Rheum.* 48(8), 2132–2145 (2003)
16. Alizadeh, A., et al.: Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403(4), 503–511 (2000)
17. van der Pouw Kraan, T.C.T.M., Wijbrands, C.A., van Baarsen, L.G.M., Voskuyl, A.E., Rustenburg, F., Baggen, J.M., Ibrahim, S.M., Fero, M., Dijkmans, B.A.C., Talk, P.P., Verweij, C.L.: Rheumatoid arthritis subtypes identified by genomic profiling of peripheral blood cells: Assignment of a type I interferon signature in a subpopulation of patients. *Ann. Rheum. Dis.* 66(8), 1008–1014 (2007)
18. Welsh, J.B., Sapinoso, L.M., Su, A.I., Kern, S.G., Wang-Rodriguez, J., Moskaluk, C.A.: Analysis of gene expression identifies candidate markers and pharmacological targets in prostate Cancer. *Cancer Res.* 61, 5974–5978 (2001)

Top Down Hierarchical Histogram Based Approach for Printed Devnagri Script Charecter Isolation

Trusha Gajjar ¹, Rekha Teraiya ², Gunvantsinh Gohil ³, and Mahesh Goyani ⁴

¹Department of Computer Engineering, ICCT, Gujarat Technological Uni, Anand, India

²Department of CSE, GEC, Gujarat Technological Uni, Gandhinagar, India

³Department of Computer Engineering, LDRP, Gujarat Technological Uni, Gujarat, India

⁴Department of Computer Engineering, LDCE, Gujarat Technological Uni, Ahmedabad, India

trusha_comp@yahoo.com, rekha.teraiya@live.in,

{gunvantshinh,mgoyani}@gmail.com

Abstract. According to recent survey, there are at least 550 million people are using Devnagari script for communication. Hindi is one of the language, which is derived from Devnagari script. As it is a national language of India, Hindi Optical Character Recognition (OCR) System has a wide application in areas like post offices, Library automation, License Plate Recognition, Defense organization and many more government sectors. Research on printed and handwritten character recognition has been started before 50 years. Most of the research has been done for European texts. For any character recognition system, essential step is to identify individual character and find features to compare it with the template features. In this paper, we have proposed histogram based hierarchical approach for isolating individual character from the image document. For recognizing the characters, we can use Euclidean distanc, Pearson coefficient, chain code etc. Result shows that our system is quite robust and provides accuracy up to 92% for the charecter isolation.

Keywords: Preprocessing, Segmentation, Skew correction, Histogram, Shirorekha.

1 Introduction

Optical Character Recognition (OCR) is a process by which we convert printed document or scanned page to ASCII Character that a computer can recognize [1]. OCR can be considered as an application of pattern recognition, artificial intelligence and machine vision [2]. A survey on the handwritten recognition has been carried by Plamondon et al [3], Koerich et al [4] and Arica et al [5]. A review on work done for the character recognition before 1990 is reported by Govindan et al [6]. The detail survey about the work done for Indian languages script recognition is made by Pal et al [7]. The work on machine printed Devanagari has been made by Bansal et al [8], Pal et al [9] and Chaudhuri et al [10]. The work on handwritten Devanagari numeral is carried by Hanmandlu et al [11] and Bajaj et al [12]. Some models that have been implemented for the Hand written Character Recognition system are described in [13], [14], [15], [16]. Multi font character recognition scheme suggested by Kahan

and Pavlidis [17]. Roy and, Chatterjee [18] presented a nearest neighbor classifier for Bengali characters employing features extracted by a string connectivity criterion. Abhijit Datta and Santanu Chaudhuri [19] suggested a curvature based feature extraction strategy for both printed and handwritten Bengali characters. B.B. Chaudhuri and U.Pal [20] combined primitive analysis with template matching.

Devnagari is the second most popular language in the Indian subcontinent and third most popular in the world [1], [2], [5]. Therefore recognition of Devnagari character is a special interest to us. Many works already done in this area and various strategies have been proposed by different authors. Devnagari script is in two formats, printed and handwritten. The Devnagari script has descended from the Brahmi script around the 11th century AD. Hindi is a direct descendant of Sanskrit through Prakrit and Apabhramsha [2], [5], [21]. It is the world's third most commonly used language after Chinese and English. Thus, research on Devnagari script, mainly the Hindi language, attracts a lot of interest.

Scanned document is preprocessed and segmented in lines, words and characters in top down manner using histogram, which is the least complex implementation. Accurate OCR system speed ups the procedure with decreased possible human errors [2], [5], [22]. In OCR system, lines and words are identified from the Devnagari script. Identified words are then segmented into individual characters. Post processing is applied to improve the performance. Systems for Indian scripts, as well as many low density languages are still under research stage. It is due to technical challenges and because of lack of a commercial market. [2], [23]. Rest of the paper is organized as follows. Next section describes basics of principle component analysis. Section III shows implementation approach followed by experimental results and conclusions in next section.

2 Basics of Devnagari Script

Before describing what types of features can be used to identify Devnagari script words from document images, we examine the appearance of Devnagari script. The basic set of symbols of Devnagari script consists of 33 consonants (or *vyanjan*) and 13 vowels (or *swar*) as shown in following tables 1 and table 2 respectively.

Table 1. Consonants of Devnagari Script

क	ख	ग	घ	ङ	च	छ	ज	झ	ट	ठ
ड	ढ	ण	त	थ	द	ध	न	प	फ	ब
भ	म	य	र	ल	व	श	ष	स	ह	ज्ञ

Table 2. Vowels of Devnagari Script

अ	आ	इ	ई	उ	ऊ	ऋ
ए	ऐ	ओ	औ	अं	अः	

The script has a set of modifier symbols which represent the modified shapes undertaken by the vowels, when they are combined with consonants [1],[2],[5],[22],[23] as shown in table 3. These symbols are placed either on top, at the bottom, on the left, to the right or a combination of these.

Table 3. Modifiers of Devnagari Script

ॠ	ि	ी	ु	ू	ृ	ँ	ं	ो	ौ	ँ	ः
---	---	---	---	---	---	---	---	---	---	---	---

Table 4. Vowels and Corresponding Modifiers

ॠ	ि	ी	ु	ू	ृ	ँ	ं	ो	ौ	ँ	ः
आ	इ	ई	उ	ऊ	ऋ	ए	ऐ	ओ	औ	अं	अः
का	कि	की	कु	कू	कृ	के	कै	को	कौ	कं	कः

All the individual characters are joined by a head line called “*Shiro Rekha*”. There are various isolated dots, which are vowel modifiers, namely, “*Anuswar*”, “*Visarga*” and “*Chandra Bindu*”, which add up to the confusion [1],[22]. As shown in figure 1 Devnagari word is written into the three zones or strips namely: *Middle zone* (1), *Upper zone* (2), and *Lower zone* (3). The upper zone and middle zone are differentiated by the head line (A), while the lower modifier is attached to the core character [23]. There is no corresponding feature to separate the bottom zone and middle zone. For completeness of understanding, we have shown *virtual base line* (B). But it is not useful for any kind of feature extraction as it is just virtual.

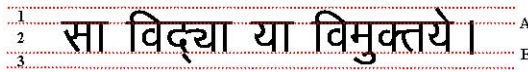


Fig. 1. Typical Writing of Devnagari Script

The upper zone contains the top modifiers, and the lower zone contains the lower modifiers. In a Hindi word, the top and bottom strips are not always necessary, but depend on the top and lower modifiers. The occurrence of a header line in a Hindi word is a powerful feature that can be used to identify Hindi words document images [1], [2], [22]. The concept of uppercase and lowercase is absent in Devnagari script and writing style of Devnagari is from left to right in a horizontal manner. Because of presence of modifiers (*matras*) in all three zone, characters like ग, ङ, ण, श etc. (which are converted in to two units when Shirerekha is removed), curve shape, compound characters etc reasons segmentation and recognition of Devnagari script is very difficult task compare to English language. Recognition of printed characters is

itself a challenging problem since there is a variation in the same character due to different font family, font size, font orientation etc. Besides, same font and size, there may also have bold face character as well as normal one. Thus, width of the stroke is also a factor that affects recognition [18]. There may be noise pixels that are introduced due to scanning of the image. Therefore, a good character recognition approach must eliminate the noise after reading binary image data; smooth the image for better recognition [1].

3 Implementation Approach

Because of different font family, sizes, orientation, weight etc it is difficult task to recognize characters [1], [18]. Character isolation is the first step towards character recognition. Figure 2 explains the complete flow of the operation. Several preprocessing steps are required to implement OCR system. Preprocessing is essential step to produce data that are easy for the OCR to operate at high robustness and accuracy. Preprocessing includes digitization, binarization, noise removal, skew detection and correction, segmentation in various levels and scaling. We will discuss over proposed approach for below text, which is shown in both, Devnagari and English.

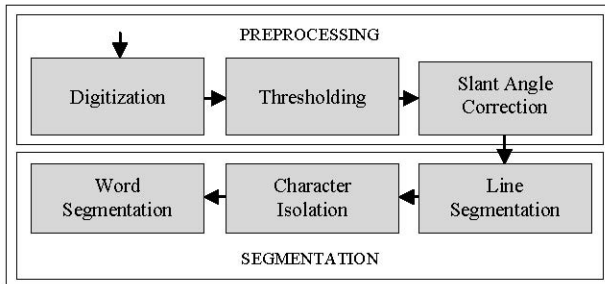


Fig. 2. Processing Steps

रात आंखो मे ढली
 पलको पे जूगनू आये
 हम हवाओ की तरह
 जा के उसे छू आये

Fig. 3. Test document in hindi

English conversion (*not translation*) of above four lines is as follow:

Raat Aankho Me Dhali, Palko Pe Jugnu Aaye
 Hum Hawao ki Tarah, Ja Ke Use Chhu Aaye

3.1 Digitization

The process of text digitization can be performed either by a scanner, computer or by a digital camera. We have used a computer generated text document. The digitized images are in gray tone.

3.2 Thresholding and Morphological Operations

Binarization is a technique by which the gray scale images are converted to binary images. Though we generate document on computer, while we save it, because of quantization, some noise is generated. This noise is appeared as gray spots on image, so it is necessary to binarize the input image. We have used a histogram-based threshold approach to convert gray scale image into two tone image. Morphological operations like opening and closing are used to remove noise and join disjoint character edge points. Figure 4 shows the histograms of original image and binary image.

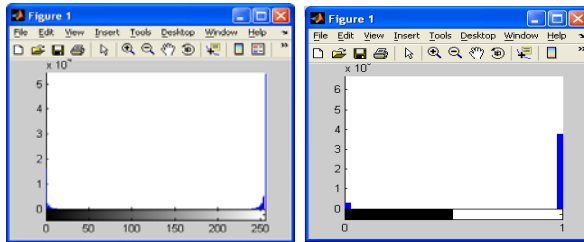


Fig. 4. Histogram of original Image (Left) and Thresholded Image (Right)

Scanned documents often contain noise that arises due to printer, scanner, print quality, age of the document, etc. Therefore, it is necessary to filter this noise before we process the image. The commonly used approach is to apply low pass filter to the image. Pepper and salt noise can be best handled with median filter. Only salt and only pepper noise can be removed effectively using min and max filters respectively.

3.3 Slant Angle Correction

When document to be scanned is fed to the scanner carelessly, digitized image may be skewed. Identification of skew angle and correction of it is essential. Skew correction can be done by rotating document in inverse direction by same skew amount. An approach based on the observation of head line of Devnagari script used for skew detection and correction. Keep rotating the document by angle θ and find out the maximum row histogram value. We will get maximum value for row histogram when the headline gets aligned with horizontal direction. After that, further rotation of document decreases the maximum value of row histogram.

3.4 Segmentation

We have employed hierarchical approach for segmentation. Segmentation is performed at different levels like line segmentation, word segmentation, character segmentation, zone wise matra Isolation. Our assumption is that intensity levels are normalized and inverted, it means white pixel indicates intensity zero and black pixel indicates intensity one.

3.4.1 Line Segmentation

Histogram enjoys the central position in segmentation. From histogram of the test image, it is very easy to calculate the boundaries of each line. For isolating text lines, image document is scanned horizontally to count number of pixels in each row. Frequency of black pixels in each row is counted in order to construct the row histogram. This count becomes zero between line gapes. Row histogram of test document of Figure 3 is shown in figure 5 for each line.

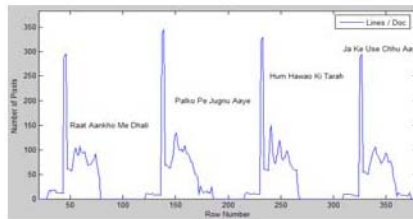


Fig. 5. Histogram for Lines per Document

Segmented lines of test document are shown in figure 6.

रात आंखो मे ढली
 पलको पे जूगनू आये
 हम हवाओ की तरह
 जा के उसे छू आये

Fig. 6. Segmented line of test document

3.4.2 Word Segmentation

Column histogram of each segmented line gives us the boundaries of each word. The portion of the line with continuous black pixels is considered to be a word in that line. If no black pixel is found in some vertical scan that is considered as the spacing between words. Thus different words in different lines are separated. Figure 7 and figure 8 shows the column histogram and actually separated words of line -1 respectively. (रात आंखो मे ढली - Raat Aankho Me Dhali)

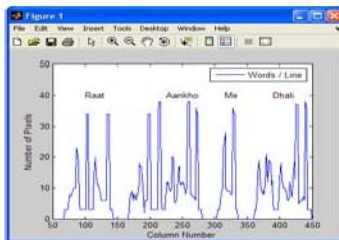


Fig. 7. Histogram for Words per Line

रात आंखो मे ढली

Fig. 8. Segmented words of Line - 1

3.4.3 Charecter Segmentation

Head Line / Shirorekha Detection: In Hindi, all characters are connected with the head line. To segment the individual character from the segmented word, we first need to find out the headline of the word which is also called 'Shirorekha'. From the word, a row histogram is constructed by counting frequency of each row in the word. The row with highest frequency value indicates the headline. Because of higher font size and weighted fonts sometimes there are consecutive two or more rows with almost same frequency value. In that case, 'Shirorekha' row is not a single row. Rather all rows that are consecutive to the highest frequency row and have frequency very close to that row constitute the Shirorekha which is now thick headline.

पलको पे जगन आय

Fig. 9. Segmented characters of line 2

Detection of character / Modifiers in Middle zone: As figure 9 shows, head line is removed first so that character segmentation becomes very easy. To find the characters from word, column histogram is calculated. Zero pixel count gives boundary of the character. Figure 10 shows the result of character segmentation of middle zone of line 1.

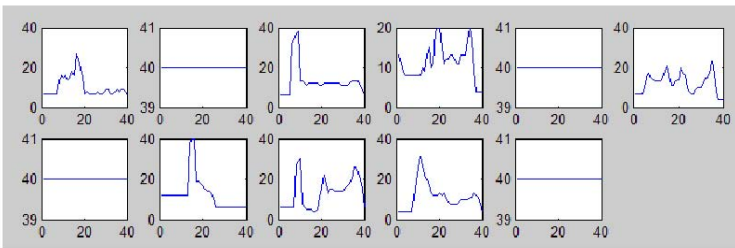


Fig. 10. Histogram for individual Characters of Middle Zone (Without Head Line)

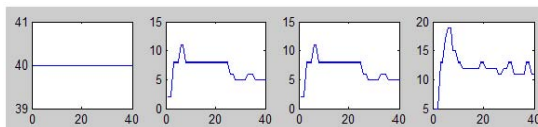


Fig. 11. Histogram Matra of Upper Zone (First Line)

Detection of Modifiers in Upper zone: To find the portion of any character above the ‘Matra’, then we can move upward from the ‘Matra’ row from a point just adjacent to the ‘Matra’ row and between the two demarcation lines. If it is, then a greedy search is initiated from that point and the whole character is found.

Detection of Modifier in Lower zone: To segment the characters below another character, baseline of the segmented word has been calculated. Like head line, there is no base line in Devnagari script, so isolation of matra in lower zone is bit hard compare to isolation of matra in upper zone. After careful investigation, we came to the conclusion that row histogram of the section below head line looks like figure 12(a), if no lower zone modifier is available and is look like figure 12(b) if lower zone modifiers are present. Histogram statistics provides very simple estimation of base line.

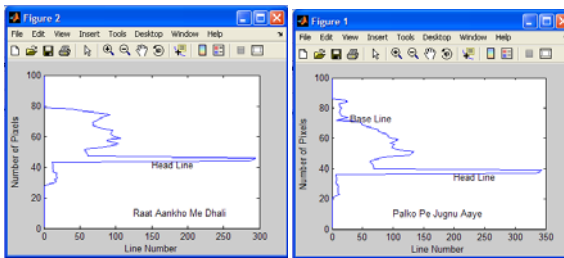


Fig. 12. (a). Row Histogram of first line, (b) Row Histogram of second line

Once base line is estimated, modifier detection job is too simple. Now the detection of matra is identical to that of in upper zone. Figure 13 illustrates the segmented two lower zone modifiers of line 2.

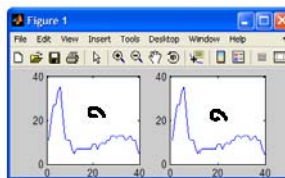


Fig. 13. Histogram Matra of Lower Zone (Second Line)

4 Experimental Results

Discussed approach is very robust in detection of individual characters. Beauty of this approach is that it is very simple and computationally also very chip as it involves histogram calculation only. Bellow chart shows detection rate of various characters and modifiers in three test documents.

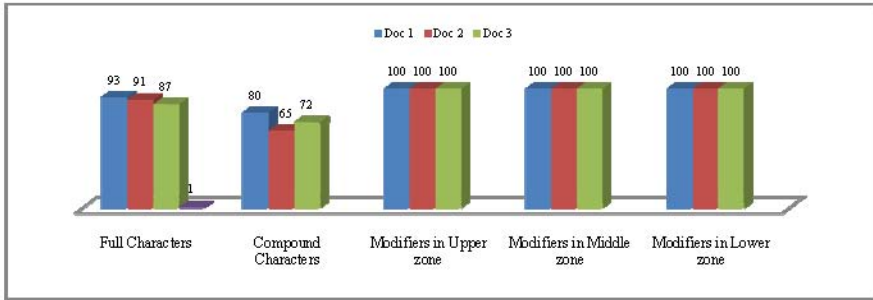


Fig. 14. Result Analysis

5 Conclusion

Experimental results show that proposed scheme is giving quite acceptable results for all the characters and modifiers accept compund characters. Post processing can be applied to improve the result of algorithm. As the proposed algorithm utilizes only histogram based apporach, its very easy to implement. It is quite faster as there is no complex processing involved. This approach is for uni font, in future it can be extended for multi font system too. We can apply PCA or LDA for feature extraction. We can use artificial neural network or support vector machine kind of classifiers for the recognition process.

References

1. Singh, R., Yadav, C.S., Verma, P., Yadav, V.: Optical Character Recognition for Printed Devnagari Script Using Artificial Neural Network. *IJCS* 1(1), 91–95 (2010)
2. Alam, M.M., Abul Kashem, M.: A Complete Devnagari OCR System for Printed Charecters. *JCIT* 1(1) (2010)
3. Plamondon, R., Srihari, S.N.: On-line and Off-line Handwriting Recognition: a Comprehensive Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(1), 63–84 (2000)
4. Koerich, A.L., Sabourin, R., Suen, C.Y.: Large Vocabulary Off-line Handwriting Recognition: a Survey. *Pattern Analysis Applications* 6, 97–121 (2003)
5. Arica, N., Yarman-Vural, T.: An Overview of Character Recognition Focused on Off-line andwriting. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews* 31(2) (2001)
6. Govindan, V.K., Shivaprasad, A.P.: Character Recognition – a Review. *Pattern Recognition* 23(7) (1990)
7. Pal, U., Chaudhuri, B.B.: Indian Script Character Recognition: A Survey. *Pattern Recognition* 37, 1887–1899 (2004)
8. Bansal, V.: Integrating Knowledge Sources in Devanagari Text Recognition System, Ph. D Thesis (1996)
9. Pal, U., Chaudhuri, B.B.: Printed Devanagari Script OCR system. *Vivek* 10, 12–24 (1997)
10. Chaudhuri, B.B., Pal, U.: An OCR System to Read Two Indian Language Scripts: Bangla and Devanagari. In: *Proceedings of International Conference on Document Analysis and Recognition*, pp. 1011–1015 (1997)

11. Hanmandlu, M., Ramana Murthy, O.V.: Fuzzy Model Based Recognition of Handwritten Hindi Numerals. *Pattern Recognition* 40(6), 1840–1854 (2006)
12. Bajaj, R., Dey, L., Chaudhury, S.: Devanagari Numeral Recognition by Combining Decision of Multiple Connectionist Classifiers. *Sadhna* 27, Part 1, 59–72 (2002)
13. Chan, K.F., Yeung, D.Y.: Elastic structural mapping for online handwritten alphanumeric character recognition. In: *Proceedings of 14th International Conference on Pattern Recognition*, Brisbane, Australia, pp. 1508–1511 (August 1998)
14. Li, X., Plamondon, R., Parizeau, M.: Model-based online handwritten digit recognition. In: *Proceedings of 14th International Conference On Pattern Recognition*, pp. 1134–1136, Brisbane, Australia (August 1998)
15. Manke, S., Bodenhausen, U.: A connectionist recognizer for online cursive handwriting recognition. In: *Proceedings of ICASSP 1994*, vol. 2, pp. 633–636 (1994)
16. Schomaker, L.R.B., Teulings, H.L.: A handwriting recognition system based on the properties and architectures of the human motor system. In: *Proceedings of the IWFHR, CENPARMI, Concordia, Montreal*, pp. 195–211 (1990)
17. Kahan, S., Pavlidis, T.: Recognition of printed characters of any font and size. *IEEE Trans. Pattern Anal. and Mach. Inte.* (9), 274–288 (1987)
18. Roy, A.K., Chatterjee, B.: Design of nearest neighbor classifier for Bengali character recognition. *J. IEEE* 30 (1984)
19. Dutta, A., Chaudhury, S.: Bengali Alpha- Numeric Character Recognition Using Curvature Features. *Pattern Recognition* 26, 1707–1720 (1993)
20. Chaudhuri, B.B., Pal, U.: A Completc Printed Bangla OCR System. *Pattern Recognition* 31, 531–549 (1997)
21. Arora, S., Bhattacharjee, D., Nasipuri, M., Malik, L., Kundu, M., Basu, D.K.: Performance Comparison of SVM and ANN for Handwritten Devnagari Character Recognition. *IJCSI* 7(3) (May 6, 2010)
22. Holambe, A.N., Thool, R.C., Jagade, S.M.: Printed and Handwritten Character and Number Recognition of Devnagari Script Using Gradient Features. *International Journal of Computer Applications* 2(9), 38–41 (2010)
23. Kumar, V., Sengar, P.K.: Segmentation of Printed text in Devnagari Script and Gurumukhi Script. *International Journal of Computer Application* 3(8) (June 2010)

Unsupervised Medical Image Classification Based on Skew Gaussian Mixture Model and Hierarchical Clustering Algorithm

Nagesh Vadaparathi¹, Srinivas Yarramalle², and Suresh Varma. P.³

¹ Dept. of I.T., M.V.G.R. College of Engineering, Vizianagaram, India
itsnageshv@gmail.com

² Dept. of IT, GITAM University, Visakhapatnam, India
sriteja.y@gmail.com

³ Aadikavi Nannaya University, Rajahmundry, India
vermaps@yahoo.com

Abstract. A novel segmentation algorithm for brain images is proposed using finite skew Gaussian mixture model. Recently, much work has been reported in medical image segmentation. Among these techniques, finite Gaussian mixture models are considered to be more recent and accurate. However, in this approach, a number of segments that an image can be divided are taken through apriori and if these segments are not initiated properly it leads to misclassification. Hence, to overcome this disadvantage, we proposed an algorithm for Medical Image Segmentation using Hierarchical Clustering and Skew Gaussian Mixture. The experimentation is done with four different brain images and the results obtained are evaluated using Quality metrics.

Keywords: Segmentation, finite Gaussian mixture model, Skew Gaussian distribution, Classification, image quality metrics.

1 Introduction

The research in medical field has geared up tremendously in the recent years. This may be due to the fact that many new diseases related to brain were emerging which needs rigorous research to trigger these diseases. As the research in this direction is spreading, many new models are available in literature to identify these diseases. Among these, model based MRI brain segmentation has gained popularity [1-5]. The brain is surrounded with many other tissues apart from White Matter (WM), Gray Matter (GM) and Cerebro Spinal Fluid (CSF), in addition some noise gets embedded into the system as a default at the time of acquisition and inhomogeneity in magnetic fields also aid to considerable changes in medical images.

Some diseases that are related with brain are Parkinson's, Acoustic neuroma, which leads to memory loss and hearing loss. Effective segmentation techniques help to identify these diseases there by driving towards effective treatment. Segmentation is a process of converting inhomogeneous data into homogeneous data. There are

many segmentation algorithms in literature consisting of both parametric and non-parametric models. Among these models, parametric modeling, in particular, model based on finite Gaussian mixture model has gained popularity [2-3]. This is due to the fact that, of the basic assumption that every image considered in reality follow a bell shaped distribution [6]. But the brain structures are non-rigid, complex in shape and vary from person to person [7]. Segmenting brain images is a challenging task and Gaussian mixture model models are not well suited due to the different structures of the brain. Hence, it is necessary to develop new algorithms which help in segmenting brain images more efficiently and effectively.

Hence, in this paper Skew Gaussian mixture model is proposed to cater the non-uniform patterns of the brain structure. It also includes Gaussian Mixture model as a limiting case. This paper is organized as follows: In section – 2, Hierarchical clustering is briefed and Section – 3 describes about the developed model. In section – 4, the initialization parameters are given and in section – 5, the segmentation algorithm is presented. Section – 6 deals with the experimentation that is carried out by using four types of brain images and the performance evaluation that is carried out is discussed in section – 7 and in section – 8 conclusions are presented.

2 Hierarchical Clustering

Clustering aims at partitioning the data without using the training data, hence, they are called unsupervised models. Clustering is defined as a technique where the objects of interest with similarity along the dimension of interest are kept close and the other objects are apart. The dimension of interest depends on the application [8].

A Hierarchical Clustering goes one step further by collecting similar clusters at different levels into a single cluster by forming a tree which gives better selection of clusters for further exploration and hence, in this method Hierarchical Clustering is utilized.

Given a set of N items to be segmented and an $M \times N$ distance (or similarity) matrix, the basic process of hierarchical segmenting is as follows.

- (1) First, assign each item to a segment, so that if we have N items, it implies that we have N segments, each containing just one item. Let the distances (similarities) between the segments be the same as those (similarities) between the items they contain.
- (2) Find the closest (most similar) pair of segments and merge them into a single segment, i.e. we will now have one segment less.
- (3) Compute distances (similarities) between the new segment and each of the old segments.
- (4) Repeat steps 2 and 3 until all items are segmented into a single segment of size N .

Step 3 can be done using single-linkage method. In single-linkage segmenting (also called the connectedness or minimum method), we consider the between one segment and another to be equal to the shortest distance from any member of one segment to any member of the other segment. If the data consist of similarities, we consider the

similarity between one segment and another to be equal to the greatest similarity from any member of one segment to any member of the other segment. The $M \times N$ proximity matrix is $D = [d(i, j)]$. The segmenting is assigned sequence numbers $0, 1, \dots, (n - 1)$ and $L(k)$ is the level of the k^{th} segmenting. A segment with sequence number m is denoted as (m) and the proximity between segments (r) and (s) is denoted as $d[(r), (s)]$. The algorithm is composed of the following steps:

- (1) Start with the disjoint segments having level $L(0) = 0$ and sequence number $m = 0$.
- (2) Find the least dissimilar pair of segments in the current s , say pair $(r), (s)$, where the minimum is over all pairs of segments in the current segmenting.
- (3) Increment the sequence number: $m = m + 1$. Merge segments (r) and (s) into a single segment to form the next segmenting m . Set the level of this segmenting to $L(m) = d[(r), (s)]$.
- (4) Update the proximity matrix, D , by deleting the rows and columns corresponding to segments (r) and (s) and adding a row and column corresponding to the newly formed segment. The proximity between the new segment, denoted (r, s) and the old segment (k) is defined as $d[(k), (r, s)] = \min(d[(k), (r)], d[(k), (s)])$.
- (5) If all objects are in one segment, stop. Else, go to step.2

3 Skew Gaussian Distribution

The pixels intensities inside the medical images may not be symmetric or bell shaped due to several factors associated like part of the body, bone structure etc. In these cases, the pixels are distributed asymmetrically and follow a skew distribution. Hence, to categorize these sorts of medical images, Skew Gaussian distribution is well suited. Every image is a collection of several regions. To model the pixel intensities inside these image regions, we assume that the pixels in each region follow a Skew normal distribution, where the probability density function is given by

$$f(z) = 2 \cdot \phi(z) \cdot \Phi(\alpha z); \quad -\infty < z < \infty. \quad (1)$$

$$\text{where, } \Phi(\alpha z) = \int_{-\infty}^{\alpha z} \phi(t) dt \quad (2)$$

$$\text{and, } \phi(z) = \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}}. \quad (3)$$

$$\text{Let, } y = \mu + \sigma z.$$

$$z = \frac{y - \mu}{\sigma} \quad (4)$$

Substituting equations 3.2, 3.3, and 3.4 in equation 3.1,

$$f(z) = \sqrt{\frac{2}{\pi}} \cdot e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \left[\int_{-\infty}^{\alpha\left(\frac{y-\mu}{\sigma}\right)} \frac{e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}}{\sqrt{2\pi}} dt \right]. \quad (5)$$

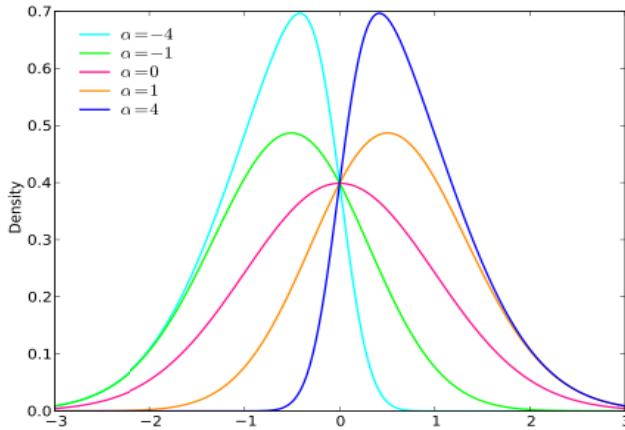


Fig. 1. Skew Normal Distributions

4 Initialization of Parameters

In order to initialize the parameters, it is needed to obtain the initial values of the model distribution. The initial estimates of the Mixture model μ_i , σ_i , λ_i and α_i where $i=1, 2, \dots, k$ are estimated using Hierarchical Clustering algorithm as proposed in section-2. It is assumed that the pixel intensities of the entire image is segmented into a K component model π_i , $i=1, 2, \dots, k$ with the assumption that $\pi_i = 1/k$ where k is the value obtained from Hierarchical Clustering algorithm discussed in section-2.

5 Segmentation Algorithm

After obtaining the initial estimates, the next step is image reconstruction by allocating the pixels to the segmentation. This operation is done by segmentation algorithm. This segmentation algorithm is given as follows:

- Step 1: Obtain the pixel intensities of the gray image. Let they be represented by x_{ij} .
- Step 2: Obtain the number of regions by k -means algorithm and divide the (image) pixel into regions.
- Step 3: For each region obtain the initial estimates using moment methods of estimation for μ_i , σ_i . Let $\alpha_i=1/k$ is the initial estimate for α_i .
- Step 4: Implement the segmentation and retrieval algorithm by considering maximum likelihood estimate.
- Step 5: With the step 4 obtain the image quality metric.

6 Experimentation

The above developed method is applied on 4 different brain images obtained from brain web images. The input images and their corresponding histograms are presented

in Figure-2a, 2b, 2c, 2d. To evaluate our algorithm we have used both T_1 and T_2 type images. In order to initialize the Hierarchical Clustering, we have used the histogram of the image as input, and basing on the peaks we have obtained the initial estimates of k . The experimentation is carried out in MATLAB environment by considering the images of fixed size. We have used both T_1 and T_2 images, where the fat decomposition varies in both these images. The segmentation process is carried out by using the segmentation algorithm presented in Section-5. The outputs images obtained after performing the segmentation are presented in figures-3a, 3b, 3c, 3d. After segmenting the performance of the segmentation results were evaluated by using Segmentation quality metrics such as Jaccard coefficient and Tanimoto. The formulas for evaluating these metrics are given below and the results obtained are tabulated and presented in Table-1.

$$\text{Jaccard Coefficient (JC)} = \frac{|X \cap Y|}{|X \cup Y|} = \frac{a}{a + b + c}$$

$$\text{Volume Similarity (VS)} = 1 - \frac{||X| - |Y||}{|X| + |Y|} = 1 - \frac{|b - c|}{2a + b + c}$$

Where, $a = |X \cap Y|$, $b = \frac{|X|}{|Y|}$, $c = \frac{|Y|}{|X|}$, $d = |\overline{X \cup Y}|$ and X, Y are input and output image intensities

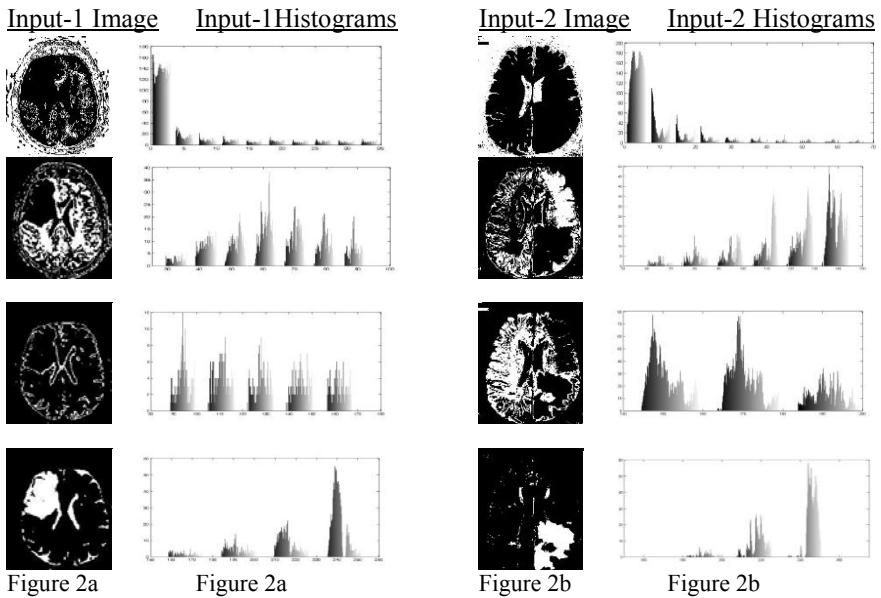


Fig. 2. Input Images with their Histograms

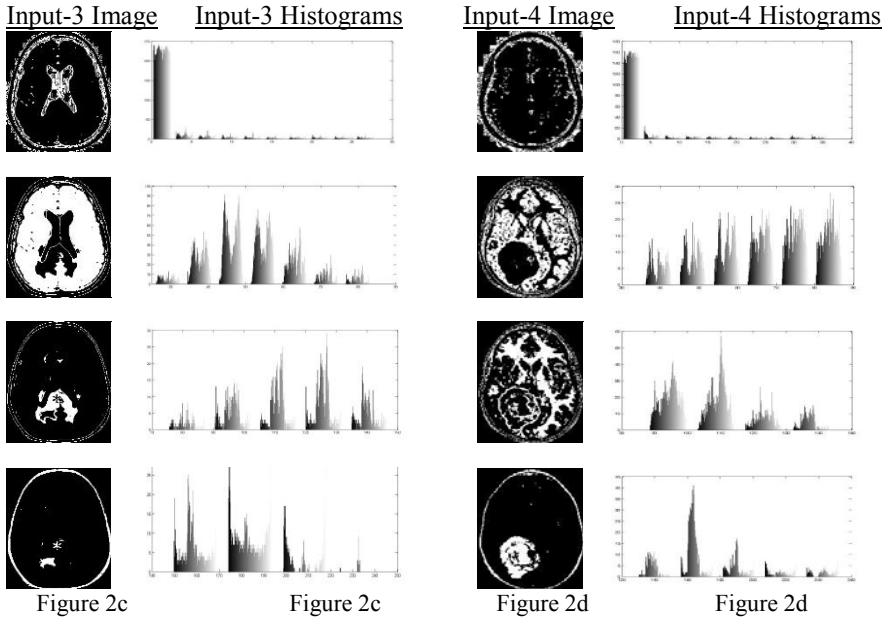


Fig. 2. (Continued)

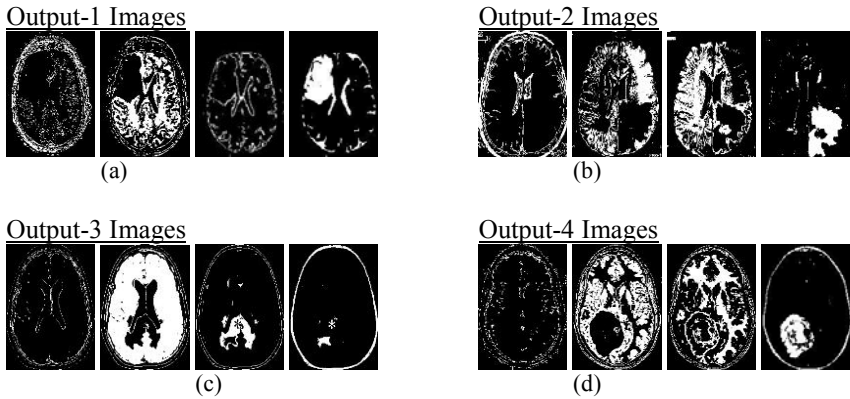


Fig. 3. Output Images

Table 1. Image Quality Metrics

Quality Metric	Values	Standard Limits
Jaccard Coefficient	0.702962	0 – 1
Volume Similarity	0.87997	0 – 1

From the above table-1, it could be easily observed that the range of segmented image is in accordance with the Segmented quality limits and hence it clearly shows that the developed model has segmented the data efficiently, since the obtained values are within the standard limits, this clearly shows the efficiency of the developed model.

7 Performance Evaluation

In order to evaluate the proposed algorithm, we have used both T_1 weighted images where water is shown as darker and fat as brighter, T_2 images where fat is shown as darker and Whiter matter is shown lighter. Among these images, T_1 images provide good gray matter and it highlights the fat decomposition. The input medical images are obtained from brain web images. We have assumed that the pixel intensities inside the brain images are non-symmetric and follow a Skew Gaussian distribution and the whole medical image is a mixture of Skew Gaussian distribution. The initialization of parameters for each segment is done using Hierarchical Clustering algorithm. The experimentation is carried out using segmentation algorithm given in section -5 and the results obtained are evaluated using Segmentation quality metrics such as Jaccard quotient (JC), Volume Similarity (VS). The reconstruction is carried out by ascribing each pixel to its appropriate position by using a method of random number generation. The performance evaluation of the retrieved images can be done by subjective testing or objective testing. Objective testing is always preferred since they are based on numeric results; hence the performance of the retrieved image is carried out by using image quality metrics such as Average difference, Maximum distance, Image Fidelity, Mean Squared Error and Signal – to – Noise Ratio. The formulae for computing the above quality metrics are as follows

Table 2. Formulas for Quality Metrics

Quality metric	Formula to Evaluate
Average Difference (AD)	$\sum_{j=1}^M \sum_{k=1}^N [F(j, k) - \hat{F}(j, k)] / MN$ <p>Where M,N are image matrix rows and columns</p>
Maximum Distance (MD)	$\text{Max}\{ F(j, k) - \hat{F}(j, k) \}$
Image Fidelity (IF)	$1 - \left[\frac{\sum_{j=1}^M \sum_{k=1}^N [F(j, k) - \hat{F}(j, k)]^2}{\sum_{j=1}^M \sum_{k=1}^N [F(j, k)]^2} \right]$ <p>Where M,N are image matrix rows and columns</p>
Mean Squared error (MSE)	$\frac{1}{MN} \sum_{j=1}^M \sum_{k=1}^N [O\{F(j, k)\} - O\{\hat{F}(j, k)\}]^2 / \sum_{j=1}^M \sum_{k=1}^N [O\{F(j, k)\}]^2$ <p>Where M,N are image matrix rows and columns</p>
Signal to noise ratio (SNR)	$20 \cdot \log_{10} \left(\frac{MAX_1}{\sqrt{MSE}} \right)$ <p>Where, MAX_1 is maximum possible pixel value of image, MSE is the Mean squared error</p>

The developed algorithm is compared with K-Means algorithm and the results obtained are tabulated and presented in Table – 3.

From the above table, it can be easily seen that the developed methods outperforms the existing method based on Medical image Segmentation based on Gaussian Mixture model and Skew Gaussian Model based on K-Means algorithm.

Table 3. Evaluated Image Quality Metrics









Image	Quality Metric	GMM	Skew GMM with K-Means	Skew GMM with hierarchical clustering	Standard Limits	Standard Criteria
	Average Difference	0.573	0.773	0.812	-1 to 1	Closer to 1
	Maximum Distance	0.422	0.922	0.9325	-1 to 1	Closer to 1
	Image Fidelity	0.416	0.875	0.923	0 to 1	Closer to 1
	Mean Squared error	0.04	0.134	0.094	0 to 1	Closer to 0
	Signal to noise ratio	17.41	29.23	33.89	$-\infty$ to ∞	Possible big
	Jaccard quotient	0.089	0.689	0.703	0 to 1	Closer to 1
	Volume Similarity	0.432	0.733	0.8799	0 to 1	Closer to 1
	Average Difference	0.37	0.876	0.749	-1 to 1	Closer to 1
	Maximum Distance	0.221	0.897	0.912	-1 to 1	Closer to 1
	Image Fidelity	0.336	0.876	0.859	0 to 1	Closer to 1
	Mean Squared error	0.2404	0.211	0.2019	0 to 1	Closer to 0
	Signal to noise ratio	14.45	35.65	39.85	$-\infty$ to ∞	Possible big
	Jaccard quotient	0.0677	0.7656	0.7921	0 to 1	Closer to 1
	Volume Similarity	0.3212	0.8767	0.8814	0 to 1	Closer to 1
	Average Difference	0.456	0.76	0.81	-1 to 1	Closer to 1
	Maximum Distance	0.345	0.879	0.807	-1 to 1	Closer to 1
	Image Fidelity	0.44	0.86	0.917	0 to 1	Closer to 1
	Mean Squared error	0.22	0.23	0.2123	0 to 1	Closer to 0
	Signal to noise ratio	19.88	37.98	39.71	$-\infty$ to ∞	Possible big
	Jaccard quotient	0.043	0.6567	0.7143	0 to 1	Closer to 1
	Volume Similarity	0.123	0.812	0.916	0 to 1	Closer to 1
	Average Difference	0.231	0.473	0.4991	-1 to 1	Closer to 1
	Maximum Distance	0.224	0.977	0.971	-1 to 1	Closer to 1
	Image Fidelity	0.212	0.813	0.892	0 to 1	Closer to 1
	Mean Squared error	0.24	0.121	0.1192	0 to 1	Closer to 0
	Signal to noise ratio	21.42	33.28	37.41	$-\infty$ to ∞	Possible big
	Jaccard quotient	0.045	0.7878	0.874	0 to 1	Closer to 1
	Volume Similarity	0.223	0.3232	0.54	0 to 1	Closer to 1
	Average Difference	0.342	0.764	0.7015	-1 to 1	Closer to 1
	Maximum Distance	0.317	0.819	0.854	-1 to 1	Closer to 1
	Image Fidelity	0.391	0.812	0.876	0 to 1	Closer to 1
	Mean Squared error	0.2514	0.228	0.1759	0 to 1	Closer to 0
	Signal to noise ratio	3.241	5.514	5.68	$-\infty$ to ∞	Possible big
	Jaccard quotient	0.141	0.776	0.791	0 to 1	Closer to 1
	Volume Similarity	0.313	0.397	0.784	0 to 1	Closer to 1

Table 3. (Continued)

	Average Difference	0.21	0.3653	0.232	-1 to 1	Closer to 1
	Maximum Distance	0.21	0.892	0.912	-1 to 1	Closer to 1
	Image Fidelity	0.2134	0.787	0.791	0 to 1	Closer to 1
	Mean Squared error	0.06	0.145	0.594	0 to 1	Closer to 0
	Signal to noise ratio	13.43	49.22	20.39	$-\infty$ to ∞	Possible big
	Jaccard quotient	0.098	0.7892	0.877	0 to 1	Closer to 1
	Volume Similarity	0.043	0.878	0.881	0 to 1	Closer to 1
	Average Difference	0.3232	0.322	0.4592	-1 to 1	Closer to 1
	Maximum Distance	0.123	0.212	0.456	-1 to 1	Closer to 1
	Image Fidelity	0.233	0.897	0.923	0 to 1	Closer to 1
	Mean Squared error	0.01	0.4345	0.119	0 to 1	Closer to 0
	Signal to noise ratio	11.11	27.267	29.86	$-\infty$ to ∞	Possible big
	Jaccard quotient	0.022	0.893	0.9124	0 to 1	Closer to 1
	Volume Similarity	0.322	0.343	0.3543	0 to 1	Closer to 1
	Average Difference	0.314	0.338	0.497	-1 to 1	Closer to 1
	Maximum Distance	0.241	0.249	0.317	-1 to 1	Closer to 1
	Image Fidelity	0.293	0.683	0.791	0 to 1	Closer to 1
	Mean Squared error	0.18	0.197	0.213	0 to 1	Closer to 0
	Signal to noise ratio	21.214	78.19	99	$-\infty$ to ∞	Possible big
	Jaccard quotient	0.455	0.762	0.815	0 to 1	Closer to 1
	Volume Similarity	0.329	0.7001	0.7158	0 to 1	Closer to 1

The performance of the developed method outperforms the Gaussian method due to the fact that the segmentation process in Gaussian Mixture model always assume that the intensities inside the medical images are bell shaped. But the shape of the images is defined on the body structure. Hence, it is necessary to go for better models. Hence, this paper mainly focuses on this disadvantage.

8 Conclusion

In brain medical analysis, segmentation plays a vital role. In particular cases such as Acoustic neuroma, it is assumed that there is a possibility of hearing loss, dizziness and other symptoms related to brain. Some acoustic neuromas can be treated with surgery. Therefore, it is needed to segment the image more accurately, which helps to identify the damaged tissues to be repaired and can be corrected by surgery. Hence, in this paper, a new novel segmentation algorithm based on Skew Gaussian distribution is proposed which helps to identify the tissues more accurately. This model is well suited in particular for medical image, where the shape of the image depends on the body structure. The performance evaluation is carried out by using quality metrics. The results show that, this developed algorithm outperforms the existing algorithm.

References

1. Pham, D.L., Xu, C.Y., Prince, J.L.: A survey of current methods in medical image segmentation. *Annu. Rev. Biomed. Eng.* 2, 315–337 (2000)
2. Van Leemput, K., Maes, F., Vandeurmeulen, D., Suetens, P.: Automated model-based tissue classification of MR images of the brain. *IEEE Trans. Med. Imag.* 18(10), 897–908 (1999)

3. Dugas-Phocion, G., González Ballester, M.Á., Malandain, G., Lebrun, C., Ayache, N.: Improved EM-based tissue segmentation and partial volume effect quantification in multi-sequence brain MRI. In: Barillot, C., Haynor, D.R., Hellier, P. (eds.) MICCAI 2004. LNCS, vol. 3216, pp. 26–33. Springer, Heidelberg (2004)
4. Van Leemput, K., Maes, F., Vandermeulen, D., Suetens, P.: A unifying framework for partial volume segmentation of brain MR images. *IEEE Trans. Med. Imag.* 22(1), 105–119 (2003)
5. Prastawa, M., Bullitt, E., Ho, S., Gerig, G.: Robust estimation for brain tumor segmentation. In: Ellis, R.E., Peters, T.M. (eds.) MICCAI 2003. LNCS, vol. 2879, pp. 530–537. Springer, Heidelberg (2003)
6. Yamazaki, T.: Introduction of EM algorithm into color Image Segmentation. In: Proceedings of ICIRS 1998, pp. 368–371 (1998)
7. Yarramalle, S., Srinivas Rao, K.: Unsupervised image segmentation using finite doubly truncated Gaussian mixture model and hierarchical clustering. *Current Science*, 71–84 (2007)
8. Bhatia, S.K.: Hierarchical clustering for image databases. *IEEE Explorer* (2005)
9. Gajanayake, G.M.N.R., et al.: Comparison of standard image segmentation methods for segmentation of brain tumors from 2D MR Images. In: ICIS 2009, pp. 301–305 (2009)
10. Bouix, S., et al.: On evaluating brain tissue classifiers without a ground truth. *Journal of NeuroImaging* 36, 1207–1227 (2007)
11. Eskicioglu, A.M., et al.: Image Quality measures and their performance. *IEEE Transaction. Comm.* 43 (1993)
12. Chawla, K.S., Bora, P.K.: PMM based segmentation of Gray – scale images. *IEEE, Los Alamitos* (2009)
13. Priebe, C.E., Miller, M.I., Ratnanather, J.T.: Segmenting magnetic resonance images via hierarchical mixture modeling. *Comput. Stat. Data Anal.* 50(2), 551–567 (2006)

An Analog VLSI Implementation for Cepstral Technique for Disparity Estimation in Stereoscopic Vision

Harshit Agarwal¹, Sheena Sharma², and Chota Markan²

¹ Faculty of Engineering, Dayalbagh Educational Institute, Dayalbagh, Agra, India

² Department of Physics and Computer Science, Dayalbagh Educational Institute, Dayalbagh, Agra, India

{Harshitagarwal89, sheenasharma11, cm.markan}@gmail.com

Abstract. The stereoscopic vision algorithms for binocular vision are very popular and widely applied. Some of the algorithms are biologically inspired. Like, Cepstral filtering technique is applied on an interlaced image, the pattern similar to that which is found in layer IV of primate visual cortex. It involves Power spectrum in computation, which is square of absolute of Fast Fourier Transform (FFT), is a complicated and hardware unfriendly. We propose a new algorithm, in which Gabor filters, instead of Power Spectrum, are applied to the interlaced image in the Cepstral algorithm. This new algorithm makes it hardware friendly as it gives us the flexibility of working with modules which can be imitated in hardware. Such as building a FFT module is a tough task in analog circuit but determining gabor energy, an alternative to it, is achieved by elementary circuits. A hardware scheme has also been proposed that can be used to estimate disparity and the idea can be extended in building complex modules that can perform real time - real image operations with a handful of resources as compared to employing complex digital FPGAs and CPLDs.

Keywords: Computer Vision, Stereo Imaging, Biologically-inspired Vision.

1 Introduction

The Stereoscopic problem is one of the very old and interesting problems till date. It arises because of the horizontal separation between two eyes. The brain uses binocular disparity to extract depth information from the two-dimensional retinal images. Stereo vision or stereoscopic vision probably evolved as means of survival. With this, we can see where objects are in relation to our own bodies with much greater precision especially when those objects are moving toward or away from us in the depth dimension. In computer vision, binocular disparity refers to the same difference seen by two different cameras instead of eyes. This ability to perceive depth, known as stereo vision, or stereopsis is made possible by the difference in viewpoints of the scene when sensed by our left and right eyes. The information about depth in a scene is of great importance because it helps us navigate in a three-dimensional environment and aids us in recognising objects of interest, among other tasks.

There are many models dedicated to find the stereoscopic disparity. There have been different approaches, proposed, to deal with this problem such as- area-based

approach [1, 2 and 3], feature-based [4, 5 and 6]. There are different classes of algorithms and out of which, some are biologically inspired algorithms, such as Cepstral filter, which is operated on an interlaced image, formatted in a way suggested in Ocular Dominance Pattern [7 and 8]. Ocular Dominance Column in the Primary Visual Cortex corresponds to alternating image patches from left and right retina. Here, the left and right views taken from left and right eyes interact for the first time. So this pattern has great similarity with the algorithm. Other method is Phase [9] and Energy method [10, 11 and 12], which involve Gabor filters to decode disparity. The Gabor filters are the band pass filters [13, 14] which has both limited spatial width and finite bandwidth and whose space is similar to the receptive field profile of simple cells in primate visual cortex [14]. These all algorithms exploit only part of the methods used to reconstruct stereoscopic disparity.

In this paper, we propose an algorithm which is more close to biological functioning. Combination of even and odd Gabor filters are applied to a window, comprises of left and right patch placed next to each other to calculate the Gabor energy. Replace Power Spectrum with the calculated Gabor Energy in Cepstral algorithm and generate disparity map.

Further, we have also proposed an equivalent analog VLSI implementation of the algorithm which can be used to estimate disparity and the idea can be extended to real world images. Two stereo images, left and right image are taken and patched together and their electrically equivalent voltages are fed to the circuit. This could be done with help of CCDs or other equivalent transducers. The electrically equivalent image is convolved with a Gabor filter of appropriate wavelength (λ). We have convolved the incoming image with 7 points even and odd Gabor filter. After the convolution, power spectrum of image is calculated using Gabor filters. Following this is the logarithmic equivalent of the voltages. The log equivalents of voltages are again convolved with the gabor and again the power spectrum is determined. Observing this power spectrum we can estimate disparity. The distance between the voltage peaks in the left image and right image is the disparity.

Analog VLSI can implement algorithms by decomposing them into numerous, simple, parallel computational units. Although the computational units are simple, highly complex and non-linear behaviors can still be obtained from them, These circuits have several advantages over implementing logic on digital as the former are cheaper, faster resource efficient.

2 Cepstral Filtering Technique

Cepstral filtering is a technique, which was developed to measure echo, which is a shifted version of signal. Now the Cepstral filtering technique is widely used in signal and speech processing and extended to image processing. Cepstrum of a signal is Power Spectrum of Log of its Power Spectrum and is operated on an interlaced image. Yeshurun and Schwartz [7] claim that Cepstral filter can be implemented using a set of band pass filters similar to those found in the visual cortex. The authors append the left image with the right image. Let an interlaced image $f(x, y)$ be composed of a single columnar pair of width 'D' [7]. The data consists of an image patch $s(x, y)$ (right eye patch) and an identical patch butted against it (left eye patch).

The interlaced image can be represented as follows.

$$f(x, y) = s(x, y) * \{\delta(x, y) + \delta(x - D, y)\} \quad (1)$$

Cepstrum is defined as:

$$\text{Cepstrum} = |F\{\ln|f(x, y)|\}|^2 \quad (2)$$

Thus, we find the disparity of the patch by locating the largest delta function. The logarithm part separates the disparity signal from the image signal in the final output and makes the Cepstral filter to be non-linear [15]. The technique is unique, as it is the method, which operates on a pattern similar to Ocular Dominance Pattern showing the similarity with biological working. But the method is difficult to realize in hardware. Cepstral algorithm consists of functions, such as power spectrum (square of absolute of FFT), which is a complicated function to implement in hardware. The solution to this is to find an easy way to estimate power spectrum and a hardware approach for it is also specified.

3 Relation between Gabor Filter and FFT

Gabor filters are closely related to the Fourier transform. In fact, the complex exponential component of the filter is actually identical to the kernel of the Fourier transform [16]. Further, the Gaussian window improves the Cepstral output [1]. Hence, the Fourier Transform of the Gaussian window would be equivalent to the Gabor filtering. There is another relation between Gabor filter and FFT. Power Spectrum, which is square of absolute of FFT is closely related to Gabor Energy [17].

$$\text{Power spectrum} = (\text{Gabor Energy})^2 \quad (3)$$

Now there are two ways in which power spectrum can be estimated using Gabor filters. In our paper, we use the relation between Power Spectrum and Gabor Energy, because the computational steps are less as compared to the other relation. The absolute is part of Power Spectrum, which needs not to calculate separately if calculating the Gabor Energy. But if we use the kernel of Fourier transform, then the step would increase to calculate the absolute and squaring to estimate power spectrum.

4 New Stereo Matching Algorithm: Cepstral Filtering Approach Using Gabor Filters

Cepstral filtering is a technique, which is applied within windows which spans an ocular dominance column pair. But since FFT was a complicated function to realize in hardware. We made use of the relation between Gabor Energy and Power Spectrum, the Power Spectrum in Cepstral operator can be estimated with the Gabor filters, then the modified system would be more close to biological functioning and the proposed model was hardware friendly to realize.

The algorithm of computing Cepstral filtering technique using Gabor filters:

1. Convolve even and odd Gabor filters with the spliced image. From the real and imaginary of the convolved values thus obtained compute the magnitude, i.e. gabor energy and square it to compute power spectrum.

2. Log Transformation of vector from step 1.
3. Now repeat step 1 with spliced image vectors being replaced by vector computed from log transformation from step 2. Doing so we have computed Cepstrum from which disparity could be inferred.

5 Hardware Implementation

We have proposed a hardware scheme that can be used to estimate disparity (fig. 1). We have built an elementary unit which could be used to estimate horizontal disparity within the limit 2. The algorithm has been tested for the disparity range up to 20 and so the hardware can also be scaled up.

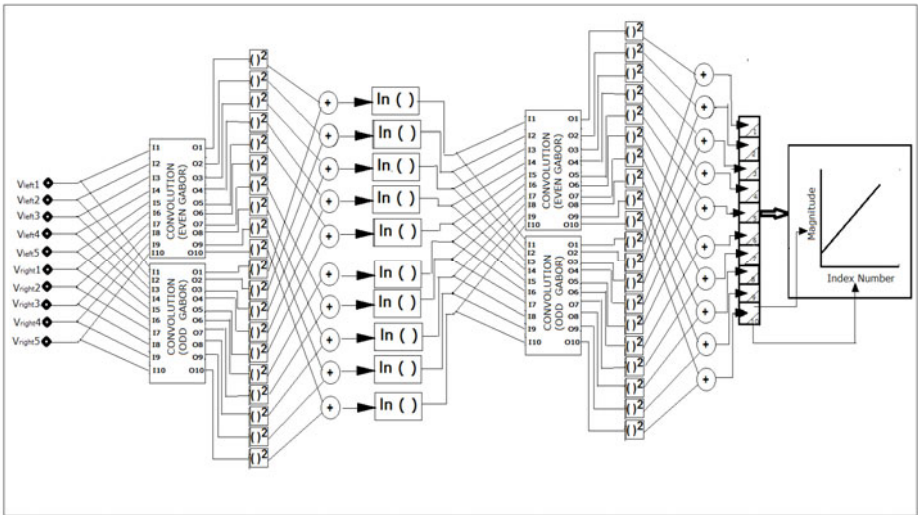


Fig. 1. A block schematic representation of the Disparity Estimation circuit

The various modules shown in the block diagram have been realized using circuits borrowed from literature [18, 19, 20 and 21] and others were self-realized.

5.1 Convolution Unit

The convolution module shown in fig. 1 is used to convolve input image with Gabor filters. The input image is basically converted into an electrical image with each pixel having a certain voltage value. Vleft1 is the voltage equivalent of pixel1 of left image and similar convention is used for other pixels. For the odd and even Gabor filters we have considered 7- discrete points [18, 19] which give equivalent 1-d even Gabor profile, fig. 2a, and 1-d odd Gabor profile, fig. 2b. We chose a disparity range -2 to 2, and accordingly our Gabor were:

$$\begin{aligned} \text{Odd Gabor} &= [-0.0498 \quad 0.4103 \quad -0.6952 \quad 0 \quad 0.6952 \quad -0.4103 \quad 0.0498] \\ \text{Even Gabor} &= [-0.0258 \quad 0.1862 \quad -0.6099 \quad 0.9059 \quad -0.6099 \quad 0.1862 \quad -0.0258] \end{aligned}$$

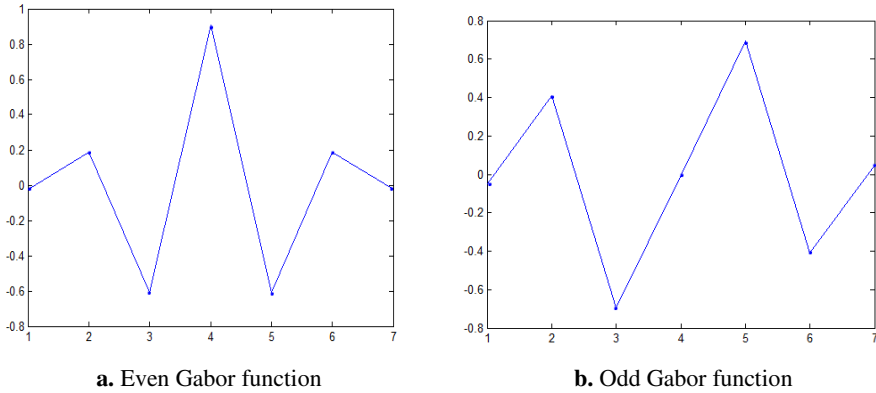


Fig. 2.

Basically convolving the image with these filters is equivalent to weighted sum of voltages and shifting the inputs and again convolving with the new weights. The weights are decided by the corresponding Gabor filters. The input conductance of the Gabor cells is made equal to the even or odd Gabor's values. The odd/even Gabor cells are realized using opamp as summing amplifier (Fig. 3 a, b).

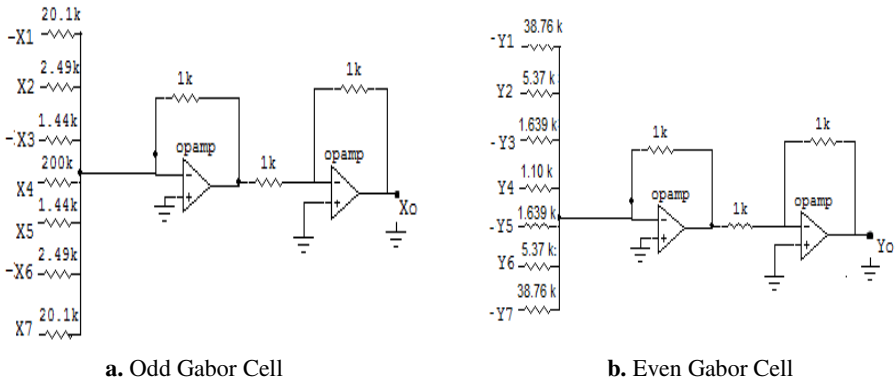


Fig. 3.

To each of the Gabor cell seven voltages are presented which computes its weighted sum.

The output voltage of odd Gabor cell (fig. 3a), X_o is

$$X_o = -0.049X_1 + 0.410X_2 - 0.695X_3 + 0.0005X_4 + 0.695X_5 - 0.410X_6 + 0.049X_7 \quad (4)$$

$X_1, X_2, X_3, X_4, X_5, X_6,$ and X_7 are the presented inputs to the Gabor cell. Negative voltages in the circuit basically refer to presenting inverted voltage.

And similarly the output voltage of even Gabor cell (fig. 3b), Y_o is

$$Y_o = -0.0258Y_1 + 0.1862Y_2 - 0.6099Y_3 + 0.9059Y_4 - 0.6099Y_5 + 0.1862Y_6 - 0.0258Y_7 \tag{5}$$

This will perform just one convolution operation so we need to shift our gabor cell over various inputs. Instead of physically shifting the gabor, we have replicated these gabor cells and they are fed by various sets of input data. And this is how we have realized a convolution unit. So these gabor cells are replicated since convolution needs shifting of inputs as well, so 10 such cells were replicated for convolution with odd gabor and 10 for even gabor filters respectively. The convolution unit is fed with inputs which are $I_1, I_2, I_3, I_4, I_6, I_7, I_8, I_9, I_{10}$, equivalent voltages of the test image. First gabor cell of the convolution block was presented with the inputs (0, 0, 0, I_1, I_2, I_3, I_4) and it generates the first output point ‘output1’ and concurrently different data sets are presented to other cells and thus in this way complete convolution is done. Here presenting ‘0’ refers to connect the input pin to ground. The scheme is depicted in fig. 4.

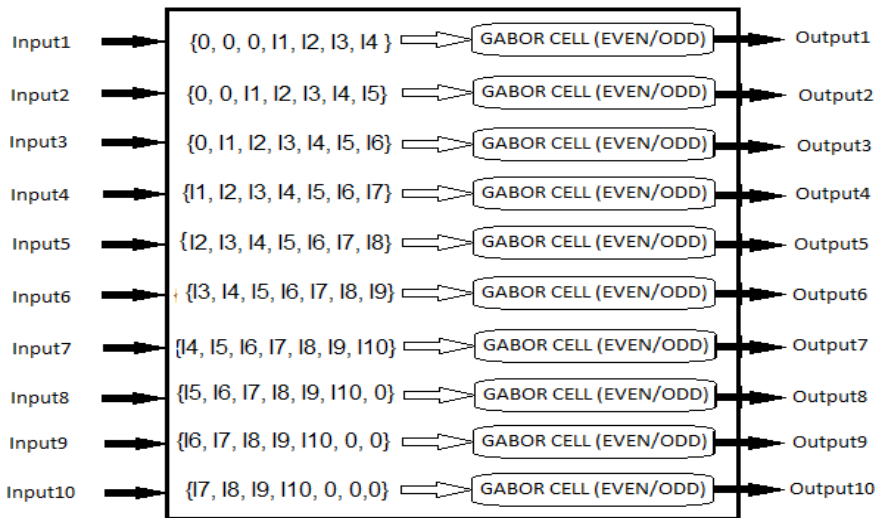


Fig. 4. Convolution Scheme

5.2 Squaring Module

Let G_c, G_s refers to resultant matrix after convolution of spliced image with even/odd Gabors respectively.

$$\text{Gabor Energy} = \sqrt{G_c^2 + G_s^2} \tag{6}$$

And from our study of algorithm as we have showed

$$\text{Power spectrum} = (\text{Gabor Energy})^2$$

Thus,

$$\text{Power spectrum} = \sqrt{(Gc^2 + Gs^2)^2} = Gc^2 + Gs^2 \tag{7}$$

For the purpose of squaring, a squaring module [20] is made use of, modified to suit our requirement as shown in fig. 5. The input voltage V_{in} is split into $+V_{in}/2$ and $-V_{in}/2$ and then fed to FVF squaring circuit [20]. The corresponding output voltage can be measured at the output node.

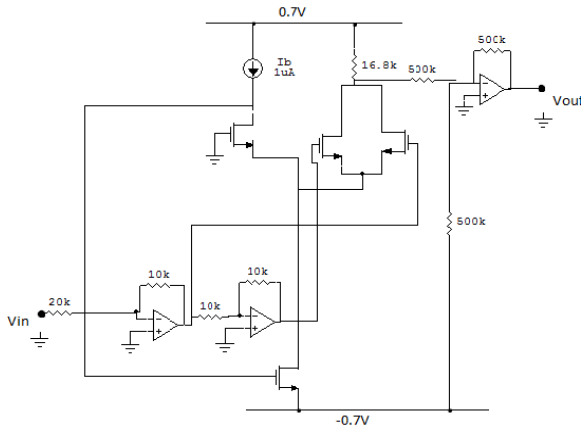


Fig. 5. FVF Squaring circuit

5.3 Summer Module

Summation was done with help of non-inverting op-amp used as a summer, as shown in fig. 6.

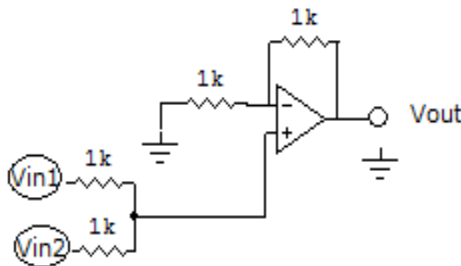


Fig. 6. Summer circuit

Output voltage of the corresponding summer is,

$$V_{out} = V_{in1} + V_{in2} \tag{8}$$

5.4 Logarithmic Converter

Once the power spectrum has been determined there is a need to compute its natural log. Though there are several circuits available for determination of log of voltages [21, 22] but in our test cases the voltages were falling in the range 50mV to 1V, so most of the circuits failed to deliver the correct result or were too complex to build. For a wider range it can be closely approximated with a higher order polynomial. Since we wanted a simple circuit with only handful of resources we estimated it with a quadratic equation.

The equivalent function, $f(x) = -3.1*(x^2) + 5.7*x - 2.8$

Both the profile have been plotted and showed in fig. 7 (blue-log profile, cyan-estimated polynomial) and fig. 8 presents a scheme to implement equivalent log circuit.

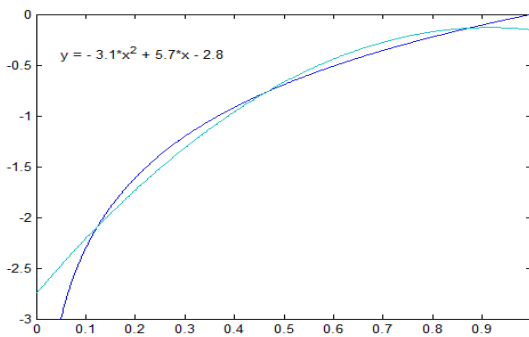


Fig. 7. Comparison of log function and estimated polynomial

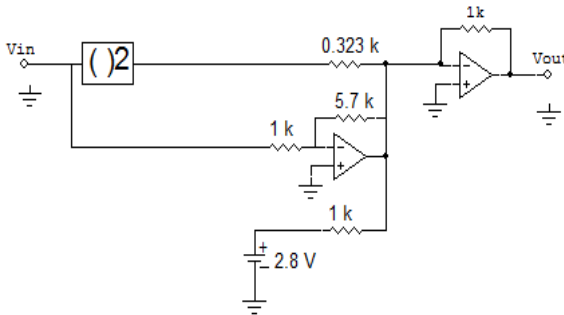


Fig. 8. Equivalent logarithmic Converter

Once the log transformation has been done, again convolve even and odd Gabor filters with the vector obtained, as shown in fig. 1. The same modules are used as a part of circuit to compute Cepstrum. From the final vector, Cepstrum, we can infer the disparity. Cepstrum algorithm states that half the distance between the two peaks is the disparity. In our case the graph plotted between the output vector voltages and their index numbers will yield disparity. Disparity's value is floor of half the difference between the index numbers of two peaks.

6 Experiment and Results

We performed our operations on a test image, shown in fig. 9

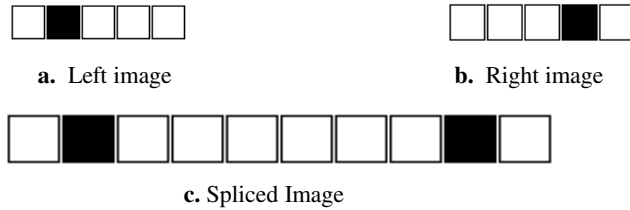


Fig. 9.

Considering each box represents a pixel and black dot is electrically equivalent to 1V and white box is 0V. Thus the Voltage equivalent of this image are 0,1,0,0,0 V for the left image, shown in fig. 9a and 0,0,0,1,0 V for the right image, shown in fig. 9b, Therefore the equivalent spliced image is (0,1,0,0,0,0,0,0,1,0) V shown in fig. 9c. Thus disparity is 2 in the given case, as the image is shifted by distance ‘2’ in horizontal direction. Such a scheme assumes that we are choosing rectified images having only horizontal disparity. To test the algorithm we simulated the algorithm on Matlab as well as the equivalent analog VLSI implementation was built and tested on Spice and from both the simulations the final vectors were observed. The observations are plotted in fig. 10 with continuous curve representing result from Matlab simulation and broken curve plotted according to spice simulation. Both Matlab and T-Spice produced logically same results.

Disparity is half the distance between the two peaks and is 2.5 from the graph. Since the disparity is always an integer value it was observed that the disparity was ‘floor’ value of the result, which was the case for other test cases as well.

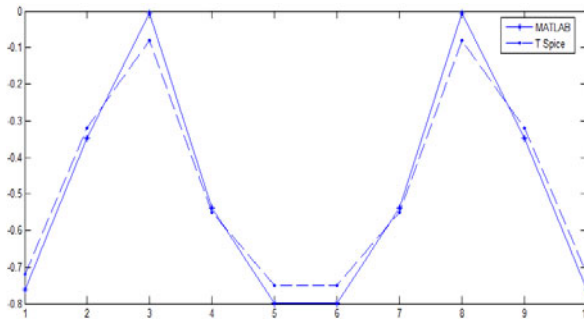


Fig. 10. Final resultant vectors in Matlab and T-Spice

We tried our circuit for another test cases having disparity within the range -2 to 2 and correct result was observed.

7 Conclusion and Future Work

The paper presents a technique for matching stereo images in the framework of Cepstral algorithm. The contribution of the paper is related to the use of Gabor filtering instead of Power Spectrum in order to make possible hardware implementation. Disparity is computed by using paired even and odd Gabor filters. The used scheme makes the algorithm close to biological functioning. The presented algorithm gives the essence of Gabor filters, which depicts the receptive fields and Ocular Dominance Pattern. Its usage of simple functions makes the algorithm suitable for hardware realization. Our work is to find out the model which depicts the brain to reconstruct the stereoscopic disparity and should be hardware-friendly. The hardware scheme has also been presented for a simple case. An analog circuit was emulated based on the algorithm. Some analog circuits were introduced in the paper such as convolution unit, simple logarithmic converter. Such sort of approach, breaking large complex systems into small elementary systems and imitating their equivalent hardware can help to build large non-linear systems which could serve the same purpose as FPGAs and other programming devices which are less resource-efficient. The idea can be scaled up to perform real time - real image operations using VLSI systems which helps in parallel programming to mimic neuro-biological architectures. As we are moving into the era of smart-pixel cameras where various operations are performed at each pixel level instead of extracting image as a whole and performing DSP operations over it, such analog implementations because of their relatively smaller sizes will play a major role. Further the improvement of the new algorithm is our current work to achieve more satisfactory results and to implement hardware for real images.

References

1. Stefano, L.D., Marchionni, M., Mattoccia, S., Neri, G.: A fast area-based stereo matching algorithm. In: 15th International Conference on Vision Interface, vol. 22, pp. 983–1005 (2002)
2. Sun, J., Shum, H.Y., Zheng, N.N.: Stereo matching using belief propagation. *Pattern Analysis And Machine Intelligence* 25, 787–800 (2003)
3. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. *Pattern Analysis and Machine Intelligence* 23(11), 1222–1239 (2001)
4. Qin, X., Li, I., Tian, S.: A new image matching algorithm with modified greedy for remote sensing imagery. In: *ISPRS Workshop on Service and Application of Spatial Data Infrastructure* (2006)
5. Gong, M., Yang, Y.-H.: Near real-time reliable stereo matching using programmable graphics hardware. In: *IEEE Computer Vision and Pattern Recognition Conference*, vol. 1, pp. 924–931 (2005)
6. Hoff, W., Ahuja, N.: Surfaces from stereo: Integrating feature matching, disparity estimation, and contour detection. *IEEE Transactions on PAMI* 11, 121–136 (1989)
7. Yeshurun, Schwartz, E.: Cepstral Filtering on a columnar Image Architecture: A fast Algorithm for binocular Stereo Segmentation. *IEEE Transaction on Pattern And Machine Intelligence* 11, 759–767 (1989)

8. Ludwig, K.O., Neumann, H., Neumann, B.: Local Stereoscopic Depth Estimation Using Ocular Stripes. In: Sandini, G. (ed.) ECCV 1992. LNCS, vol. 588, pp. 373–377. Springer, Heidelberg (1992)
9. Sanger, T.D.: Stereo disparity computation using Gabor filters. *Biol. Cybern.* 59, 405–418 (1988)
10. Ohzawa, I., DeAngelis, G.C., Freeman, R.D.: Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. *Science* 249, 1037–1041 (1990)
11. Ohzawa, I., DeAngelis, G.C., Freeman, R.D.: Encoding of binocular disparity by simple cells in the cat's visual cortex. *J. Neurophysiol.* 75, 1779–1805 (1996)
12. Ohzawa, I., DeAngelis, G.C., Freeman, R.D.: Encoding of binocular disparity by complex cells in the cat's visual cortex. *J. Neurophysiol.* 77, 2879–2909 (1997)
13. Gabor, D.: Theory of Communication. *J. Inst. Electr. Eng.* 93, 429–457 (1946)
14. Marcelja, S.: Mathematical Description of the Responses of Simple Cortical Cells. *J. Opt. Soc. Am.* 70, 1297–1300 (1980)
15. Leloglu, U.M.: Artificial Versus natural stereo depth perception. In: Halici, U., Metu (eds.) *Hints from Life to AI* (1994)
16. Maimone, M.W.: Characterizing Stereo Matching Problems using stereo Spatial Frequency, PhD thesis (1996)
17. Andrysiak, T., Choras's, M.: Image Retrieval Based On Hierarchical Gabor Filters. *Int. J. Appl. Math. Comput. Sci.* 15(4), 471–480 (2005)
18. Meneses, Y.L.: Algorithms for VLSI Stereo Vision Circuits applied to Autonomous Robots, PhD thesis (1999)
19. Marco, B.G.: Microelectronic Neural Systems: Analog VLSI for Perception and Cognition, PhD thesis (1998)
20. Sakul, C., Dejhana, K.: Squaring And Square-Root Circuits Based On Flipped Voltage Follower And Applications. *International Journal of Information Systems and Telecommunication Engineering* 1 (2010)
21. Jahromi, A.G., Abrishamifar, A., Medi, A.: A Novel Voltage- to- Voltage Logarithmic Converter with High Accuracy. *Journal of Selected Areas in Microelectronics (JSAM)* 2 (2011)
22. Schiewe, A.I., Chen, K.: Analog Logarithmic and Antilogarithmic Circuits Using Switching Transistors, *afips*, p. 121 (1957)

Finding Optimal Set of Orthogonal Polynomial Operators for Efficient Texture Feature Extraction

R. Suguna¹ and P. Anandhakumar²

¹ Research Scholar, Department of Information Technology,

² Associate Professor, Department of Information Technology,

Madras Institute of Technology, Anna University, Chennai- 600 044, Tamil Nadu, India

hitec_suguna@hotmail.com, anandh@annauniv.edu

Abstract. Feature Extraction is the method of capturing the visual content of images for indexing and retrieval. It simplifies the amount of information required to describe a large set of data. In computer vision, feature detection refers to the computation of local image features from the image. Texture is the core element in numerous computer vision applications. Orthogonal Polynomial Operators are generated from a basis operator of fixed size and their efficiency in extracting texture features is studied. These operators act as filters and their responses on images are considered as feature space. From each filtered image statistical features are extracted and an optimal operator set is designed by incorporating a feature selection approach. Mahalanobis separability metric is used in the feature selection process. The optimal operator set removes insignificant operators and thus improves the performance of texture classification. Experimental results on benchmark datasets prove the effectiveness of the proposed approach.

Keywords: Texture Classification, Orthogonal Polynomial Operator, Supervised Classification, Feature Selection.

1 Introduction

Feature detection is a low-level image processing operation. It is usually performed as the first operation on an image and examines every pixel to see if there is a feature present at that pixel. As a pre-requisite to feature detection, the input image is usually smoothed by scale-space representation and one or several feature images are computed. When feature detection is computationally expensive and there are time constraints, a higher level algorithm is used so that only certain parts of the image are searched for features. Many computer vision algorithms use feature detection as the initial step, hence, a very large number of feature detectors have been developed. These vary widely in the kinds of feature detected and the computational complexity.

Visual features can be classified as domain-specific features and general features. Domain-specific features include fingerprints, human faces and general features refer to color, texture and shape. The issue of choosing the features to be extracted should be guided by the following concerns: (i) The features should carry enough information about the image and should not require any domain-specific knowledge

for their extraction. (ii) They should be easy to compute in order for the approach to be feasible for a large image collection and rapid retrieval. (iii) They should relate well with the human perceptual characteristics since users will finally determine the suitability of the retrieved images. Once features have been detected, a local image patch around the feature is extracted. The result is known as a feature descriptor or feature vector.

Texture has an important role in natural vision, and it has been widely applied to several surface characterization problems as well. In the early years of computer vision and texture analysis Haralick et al. [1] applied texture analysis methods to remotely sensed images for doing terrain analysis. They tried to classify regions of images to predefined classes to form a description of the sensed scene. Oliver [2] used texture analysis to classify regions of SAR images to forest and non-forest classes. In addition to earth surface analysis, texture has also been utilized for characterizing the surface of more concrete objects.

Typically textures and the analysis methods related to them are divided into two main categories with different computational approaches: the stochastic and the structural methods. Structural textures are often man-made with a very regular appearance consisting of primitive patterns that are systematically located on the surface. In structural texture analysis the properties and the appearance of the textures are described with different rules that specify the kind of primitive elements present in the surface and their location. Stochastic textures are usually natural and consist of randomly distributed texture elements. The analysis of these kinds of textures is based on statistical properties of image pixels and regions. Texture analysis methods try to describe the properties of the textures in a proper way. It depends on the applications what kind of properties should be sought from the textures under inspection and how to do that. This is rarely an easy task.

Texture may be defined as the characteristic variation in intensity of a region of an image. The degrees of randomness and of regularity will be the key measure when characterizing a texture. To exploit texture in applications, the measures should be accurate in detecting different texture structures. Computational complexity should not be too high to preserve realistic use of the methods.

2 Related Work

The gray-level co-occurrence matrix approach is based on studies of the statistics of pixel intensity distributions. The early paper by Haralick et al.[3] presented 14 texture measures and these were used successfully for classification of many types of materials. However, Connors and Harlow [4] found that only five of these measures were normally used, viz. “energy”, “entropy”, “correlation”, “local homogeneity”, and “inertia”. The size of the co-occurrence matrix is high and suitable choice of d (distance) and θ (angle) has to be made.

A novel texture energy approach is presented by Laws [5]. This involved the application of simple filters to digital images. The basic filters used were common Gaussian, edge detector and Laplacian-type filters and were designed to highlight points of high “texture energy” in the image. Ade investigated the theory underlying Laws’ approach and developed a revised rationale in terms of eigen filters [6]. The

filters that give rise to low variances were taken to be relatively unimportant for texture recognition.

Recent developments include the work with automated visual inspection in work. Ojala and Pietikainen [7] proposed a multichannel approach to texture description by approximating joint occurrences of multiple features with marginal distributions, as 1-D histograms, and combining similarity scores for 1-D histograms into an aggregate similarity score. Ojala introduced a generalized approach to the gray scale and rotation invariant texture classification method based on local binary patterns [8]. Hadjidemetriou et al. [9] filtered the original image with several Gaussian filters, and constructed multiresolution histogram features, achieving reasonably good accuracy with low computing costs.

Structural and geometrical features are usually more stable in overall illumination changes than statistical features, but rely strongly on primitive detection. Huet & Mattioli [10] used basic mathematical morphology operations to detect texture primitives. Watershed segmentation based on mathematical morphology was applied for detecting textons from images by Asano et al [11]. Tuceryan and Jain proposed a method based on Voronoi tessellation to extract shape properties of textons [12].

Ahonen proved that the local binary pattern operator can be seen as a filter operator based on local derivative filters at different orientations and a special vector quantization function [13]. A framework using orthogonal polynomials for edge detection and texture analysis was presented by Ganesan [14], Krishnamoorthi [15].

3 Generation of Orthogonal Polynomial Operators

A set of orthogonal polynomials $u_0(t), u_1(t), \dots, u_{n-1}(t)$ of degrees 0, 1, 2... n-1 respectively are considered to construct the polynomial operators of different sizes. Orthogonal basis function is constructed using Gram-Schmidt process. The Gram-Schmidt process is an algorithm which removes linear dependency from a set of given vectors in an inner product space. The process removes linear dependencies from the polynomials, yielding sets of orthogonal polynomials.

The Gram-Schmidt process uses projection operator which is defined as

$$proj_u(v) = \frac{\langle v, u \rangle}{\langle u, u \rangle} u \quad (1)$$

where $\langle \mathbf{u}, \mathbf{v} \rangle$ denotes the inner product of the vectors \mathbf{u} and \mathbf{v} . This operator projects the vector \mathbf{v} orthogonally onto the vector \mathbf{u} .

The Gram-Schmitt process works as follows:

$$\begin{aligned} u_k^{(1)} &= v_k - proj_{u_1}(v_k) \text{ and} \\ u_k^{(k-2)} &= u_k^{(k-3)} - proj_{u_{k-2}}(u_k^{(k-3)}) \end{aligned} \quad (2)$$

The sequence $\mathbf{u}_1, \dots, \mathbf{u}_k$ is the required system of orthogonal vectors. For a $n \times n$ basis, n^2 operators can be generated. The set of polynomial basis operators is computed by applying the outer product over u_i and u_j where u_j is the $(i+1)$ st column vector of M where

$$|M| = \begin{pmatrix} u_0(t_1) & u_1(t_1) & \dots & u_{n-1}(t_1) \\ u_0(t_2) & u_1(t_2) & \dots & u_{n-1}(t_2) \\ \vdots & \vdots & \ddots & \vdots \\ u_0(t_n) & u_1(t_n) & \dots & u_{n-1}(t_n) \end{pmatrix} \tag{3}$$

The orthogonal basis functions for n=3 and the corresponding operator set is shown in Fig. 1.

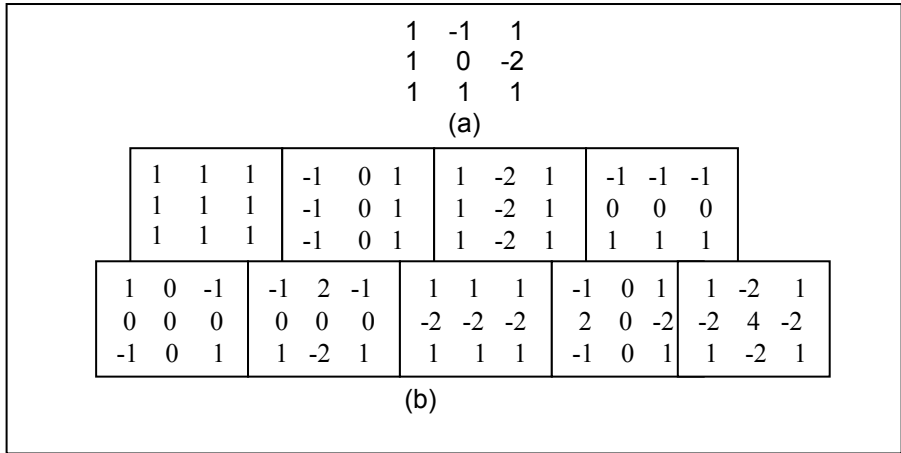


Fig. 1. (a) 3x3 orthogonal basis (b) Orthogonal Polynomial Operator Set

After the operators are generated, it is normally assumed that all the operators are useful. But only a subset of them might be useful, while others are either redundant or irrelevant. The operator selection in the design could generate a compact Orthogonal Polynomial Operator Set which also leads to low dimensional texture representations. Hence it not only saves computational cost but also tend to generalize well.

4 Selection of Optimal Operator Set

By the above analysis, feature selection is an indispensable process in the design of Optimal Operator Set. The issue is how to decide which operator should be retained and which operators can be removed. Since the ultimate goal of operators is to extract feature representations to distinguish different texture patterns, it is natural to evaluate and select operators based on the discriminative power of the features produced. An algorithm for optimal operator set selection for a supervised learning is developed.

A set of N training images from m classes were collected. Let the images be x_1, x_2, \dots, x_N . The images are convolved with the operators to form response images which constitute a feature space. Let o_1, o_2, \dots, o_p be the set of operators where p is the total number of operators. From each response image, features such as mean, standard

deviation and energy were extracted. Let f_1, f_2, \dots, f_p denotes feature group extracted from response images by the p operators, defined as

$$f_k = \begin{bmatrix} f_{1,k} \\ f_{2,k} \\ f_{3,k} \\ \vdots \\ f_{N,k} \end{bmatrix} \tag{4}$$

$f_{N,k}$ is the feature vector representing the features extracted from the Nth image response over operator k.

$$f_{N,k} = [m_{N,k} \quad s_{N,k} \quad e_{N,k}] \tag{5}$$

where $m_{N,k}$, $s_{N,k}$, $e_{N,k}$ are the mean, standard deviation and energy of the response image.

4.1 Mahalanobis Separability

To remove the correlation between the features and to describe the discriminative power among them, Mahalanobis separability has been used [16]. This measure the distance between classes i and class j along a feature dimension and removes the correlations between features. The higher the Mahalanobis separability measure, the less the correlations and the larger the discriminative power for a feature group.

The Mahalanobis separability measure J between class i and class j along feature group f_k can be defined as follows:

$$J_{i,j,k} = (M_{i,k} - M_{j,k})C_{i,j,k}^{-1}(M_{i,k} - M_{j,k})^T \tag{6}$$

where $M_{i,k}$, $M_{j,k}$ and $C_{i,j,k}$ denote mean vector and covariance matrix of training sample class i and class j in feature space along feature group f_k respectively. The feature group f_k can be evaluated based on the average Mahalanobis separability measure given by:

$$J_k = 2 X \frac{\sum_{i=1}^{m-1} \sum_{j=1}^{m-1} J_{i,j,k}}{m X (m - 1)} \tag{7}$$

Feature groups or feature group combinations with large Mahalanobis measure imply significance and should be retained. This measure is combined with forward selection algorithm to find a suitable feature subset. The selected subset of features is the representatives of the full feature set. The operators yielding this subset of features are considered as optimal operator set.

4.2 Proposed Methodology

The proposed feature selection approach has two phases.

- Construction of feature groups
- Selection of Optimal Operator Set

The details of the phases are illustrated in Fig 2 and Fig. 3.

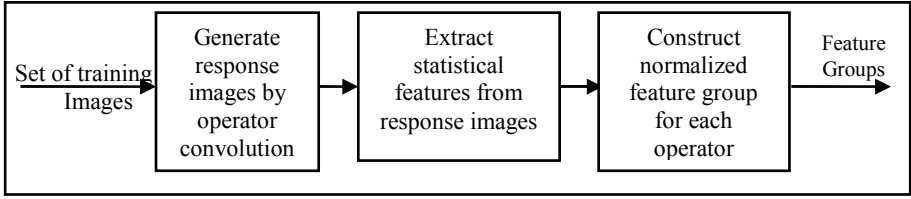


Fig. 2. Feature Group Construction

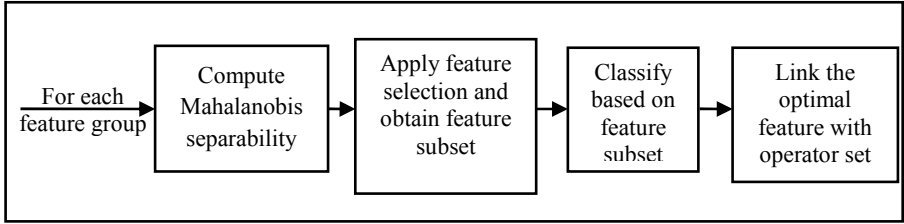


Fig. 3. Optimal Operator Set Selection

4.2.1 Feature Group Construction

Set of training images x_1, x_2, \dots, x_n are convolved with p operators and the corresponding response images forms the feature space. From the response images, statistical features are computed and the features are grouped based on the operator. Eight 3×3 operators were used and three features (mean, standard deviation and energy) were extracted from the response images. Eight feature groups were constructed and each feature group is normalized.

4.2.2 Algorithm for Selection of Optimal Operator Set

1. For each feature group f_k , calculate average Mahalanobis separability measure using Eqs (6)-(7). Feature group f_k with the maximum mahalanobis measure is selected as the first candidate feature subset and the initial subset contains that feature group $S_1 = \{f_k\}$
2. Remaining $p-1$ feature groups are combined with f_1 to form subsets S_2 of size 2. Computing the mahalanobis measures for these subsets suggest the candidate subset of size 2.
3. The above procedure for candidate feature subsets is repeated and is terminated when all p candidate feature subsets are obtained.
4. Perform classification for the training texture images using each of the above feature subsets until a candidate feature subset shows very little or no improvement in the classification performance.
5. The operators corresponding to the feature subset S_k is regarded as the final optimal operator Set.

5 Experiments and Results

The performance of the approach is demonstrated with two different problems of texture analysis. Experiment #1 is conducted on datasets to study the performance of the proposed operator set for texture classification. Experiment #2 involved a new set of images derived from Brodatz album to study the behavior of proposed approach on texture characteristics.

5.1 Experiment #1

The image data included 12 texture classes from Outex database as shown in Fig. 4. For each texture class, there were sixteen 128x128 source images. One source image from each texture class was taken and sub images are extracted by dividing the source image of size 128x128 into disjoint 32x32 images. A dataset of training images of 192 samples were generated for the experiment study. Half of the images were used for training and the remaining images were used for testing.

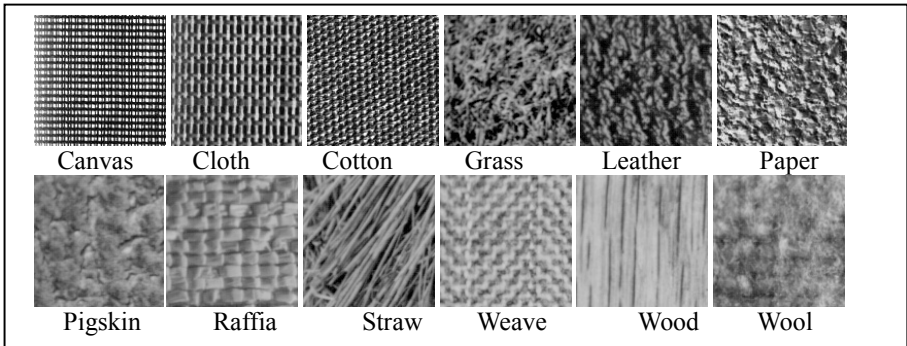


Fig. 4. Sample Images used in Experiment #1

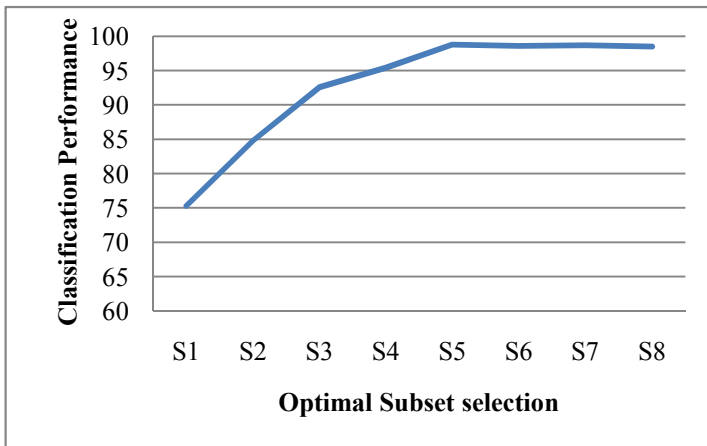


Fig. 5. Classification Performance based on candidate subset

The performance of classification is noted with original full set of operators. Then Optimal Operator Set is selected from the proposed methodology discussed in section 4.2. It is observed that only five operators are sufficient to discriminate the textures in the dataset. The classification accuracies of the feature subsets are shown in Fig. 5.

5.2 Experiment #2

The behavior of the proposed operator on different texture characteristic is studied. The image data included 12 texture images from Brodatz album which are grouped as regular and random. Sample images are shown in Fig. 6. The texture discrimination capability of proposed approach is studied on these sets. Each texture image was of size 64x64. The training dataset consisted of 144 samples (12 samples per texture class).

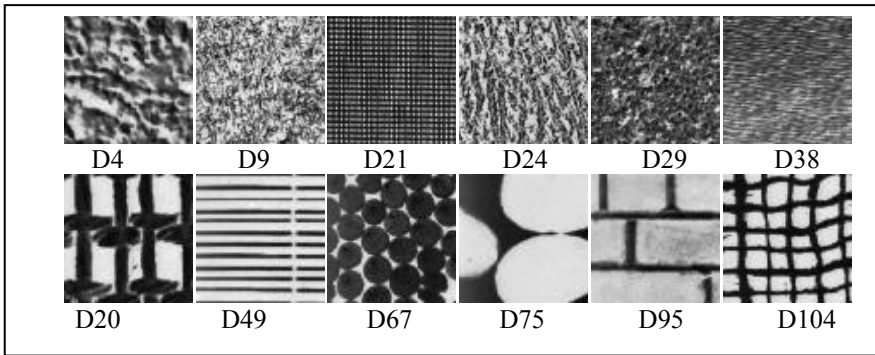


Fig. 6. Top Row with Random Texture images and Bottom Row with Regular Texture Images

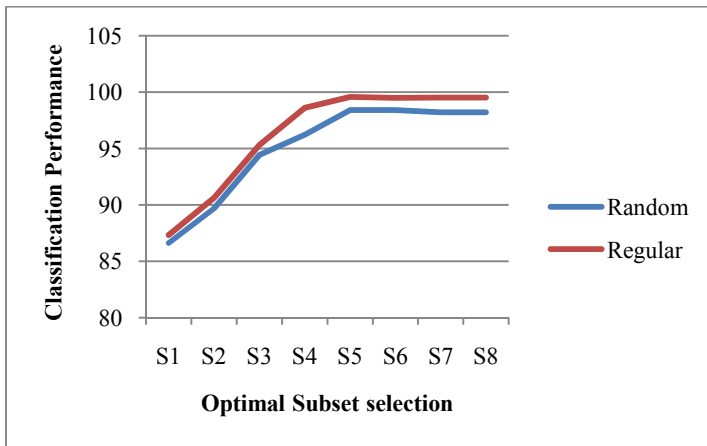


Fig. 7. Classification Performance based on candidate subset

Table 1. Comparative Study of Classification Performance with other Texture Models

Methods	Recognition rate in %
Laws Texture measure	89.2
Gabor filters	94.95
Initial Orthogonal Operator Set	95.4
Optimal Operator Set	98.7

The candidate subset performance for Regular Textures and Random Textures are shown in Fig. 7.

Table 1 shows the comparative study of classification performance with other Texture models.

6 Conclusion

Orthogonal Polynomial Operators are capable of extracting texture features from images. To improve the efficiency of classification, a methodology for finding an optimal operator set is proposed and their performance on benchmark dataset is studied. The proposed approach improves the classification performance by removing insignificant operators. This helps to save computational cost and also leads to generalization. Experimental results on benchmark datasets demonstrate the effectiveness of the approach.

References

1. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural feature for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3, 610–621 (1973)
2. Oliver, C.: Rain forest classification based on SAR texture. *IEEE Transactions on Geoscience and Remote Sensing* 38(2), 1095–1104 (2000)
3. Haralick, R.M.: Statistical and structural approaches to Texture. *Proc. IEEE* 67(5), 786–804 (1979)
4. Connors, R., Harlow, C.: A theoretical comparison of texture algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2(3), 204–222 (1980)
5. Laws, K.: Textured image segmentation. Ph.D. thesis, University of Southern California, Los Angeles, USA (1980)
6. Ade, F.: Characterization of texture by eigenfilters. *Signal Processing* 5(5), 451–457 (1983)
7. Ojala, T., Pietikainen, M.: Nonparametric multichannel texture description with simple spatial operators. In: *Proc. 14th International Conference on Pattern Recognition*, Brisbane, Australia, pp. 1052–1056 (1998)
8. Ojala, T., Pietikainen, M., Mäenpää, T.: A Generalized Local Binary Pattern Operator for Multiresolution Gray Scale and Rotation Invariant Texture Classification. In: Singh, S., Murshed, N., Kropatsch, W.G. (eds.) *ICAPR 2001*. LNCS, vol. 2013, pp. 397–406. Springer, Heidelberg (2001)

9. Hadjidemetriou, E., Grossberg, M., Nayar, S.: Multiresolution histograms and their use for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(7), 831–847 (2004)
10. Huet, F., Mattioli, J.: A textural analysis by mathematical morphology transformations: structural opening and top-hat. In: *Proc. International Conference on Image Processing*, vol. 3, pp. 49–52 (1996)
11. Asano, A., Endo, J., Muraki, C.: Multiprimitive texture analysis using cluster analysis and size density function. In: *Proc. International Symposium on Mathematical Morphology VI* (2002)
12. Tuceryan, M., Jain, A.: Texture segmentation using voronoi polygons. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12(2), 211–216 (1990)
13. Ahonen, T., Pietikainen, M.: A framework for analyzing texture descriptors. In: *Proc. Third International Conference on Computer Vision Theory and Applications (VISAPP 2008)*, Madeira, Portugal, vol. 1, pp. 507–512 (2008)
14. Ganesan, L., Bhattacharyya, P.: Edge detection in untextured and textured images—a common computational framework. *IEEE Trans. Sys. Man Cybern. B.* 27(5), 823–834 (1997)
15. Krishnamoorthi, R., Kannan, N.: A new integer image coding technique based on orthogonal polynomials. *Image and Vision Computing* 27(8) (2009)
16. Li, W., Mao, K., Zhang, H., Chai, T.: Designing Compact Gabor Filter Banks for Efficient Texture Feature Extraction. In: *Intl. Conference Control, Automation, Robotics and Vision*, pp. 1193–1197 (2010)

Design of Optimized Neuro-Wavelet Based Hybrid Model for Image Compression

Deepak Gambhir^{1,*}, Navin Rajpal², and Vipula Singh³

¹ Asst.Prof. ECE, Amity School of Engineering and Technology,
Bijwasan New Delhi
deepak@gmail.com

² Professor USIT, Guru Gobind Singh Inderprastha University,
Dwarka New Delhi

³ Professor ECE, RNS Institute of Technology, Bangalore

Abstract. Images are in its standard canonical form a matrix have significant amount of redundant data. Thus image compression methods always under wide attention for efficient multimedia data transmission and storage. This paper concerned with the design of an optimized hybrid Neuro-Wavelet based model for image compression. In this design first the images are decomposed to various sub-band via wavelet transform and then they are fed to different supervised Neural Networks which are optimized with Linear Programming. The simulation results show the clear improvement over the existing methods objectively by PSNR and subjectively by visual appearance.

Keywords: Image Compression, Back Propagation, Neural Network, Wavelet Transform, Linear Programming.

1 Introduction

Image Compression addresses the problem of reducing the amount of data required to represent an image. Image compression removes redundancy present in data pixels in such a way that reconstruction is possible. The artificial neural network has been widely used to implement a data compression after D.E.Remelhart proposed back-propagating (BP) algorithm in [8]. Some early contributions have proposed the application of neural network models to data compression in image coding systems [7]. In the direction of improving the compression rate and the quality of reconstructed image, two different variations to the basic approach have been devised and experimented: the first, change the structure of the neural networks, such as A.Namphoi et al. using multi-layer neural network to enhance the compression rate in [4]; The second, optimize the BP algorithm to improve the quality of compression [11]. However, the texture structure of image and the human visual features were considered few in their contributions, and the quality of reconstructed image is improved with sacrifice of compression rate. In this paper we describe an Image compression model uses hybrid Neuro-Wavelet model in which NN are optimized using the Linear Programming. Here,

* Corresponding Author.

Neural Network, weights & biases are updated regularly as in backpropagation learning which are optimized with linear programming concept to achieve image compression.

This paper is divided to various sections as: In section 2 we describe about the basics for wavelet transform for multiresolution analysis. In section 3 there is the explanation about the multilayer perceptron model (MLP) with back propagation learning for image compression. In section 4 the description of Linear Programming is there to solve optimum solution for the system of linear solutions. In section 5 the proposed method, in section 6 the computer simulation result and in section 7 the conclusion is drawn.

2 Discrete Wavelet Transform and Sub-Band Coding

Wavelet transform (WT) [5] represents an image as a sum of wavelet functions (wavelets) with different locations and scales. Any decomposition of an image into wavelets involves a pair of waveforms: one to represent the high frequencies corresponding to the detailed parts of an image (wavelet function) and one for the low frequencies or smooth parts of an image (scaling function). The result of wavelet transform is a set of wavelet coefficients, which measure the contribution of the wavelets at these locations and scales.

One of the big discoveries for wavelet analysis was that perfect reconstruction filter banks could be formed using the coefficient sequences $a_L(k)$ and $a_H(k)$ (Fig. 1). The input sequence x is convolved with high-pass (HPF) and low-pass (LPF) filters $a_H(k)$ and $a_L(k)$ and each result is down sampled by two, yielding the transform signals x_H and x_L . The signal is reconstructed through up sampling by two and convolution with high and low synthesis filters $s_H(k)$ and $s_L(k)$. For properly designed filters, the signal is reconstructed exactly. By cascading the analysis filter bank with itself a number of times, digital signal decomposition with dyadic frequency scaling known as discrete wavelet transform (DWT) can be formed. The mathematical manipulation that calls for synthesis is known as inverse DWT. DWT for an image as a 2-D signal can be derived from 1-D DWT.

The easiest way to obtain scaling and wavelet function for two dimensions is by multiplying two 1-D functions. The scaling function for 2-D DWT can be obtained by multiplying two 1-D scaling functions: $\psi(x, y) = \psi(x)\psi(y)$. Wavelet functions for 2-D DWT can be obtained by multiplying two wavelet functions or wavelet and scaling function for 1-D analysis. For the 2-D case, there exist three

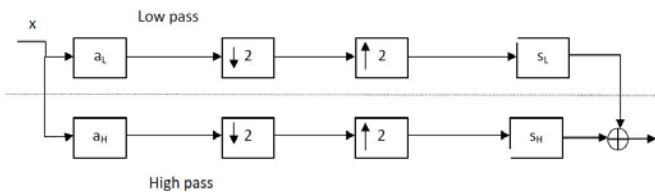


Fig. 1. Two channel Analysis and Reconstruction Filter bank

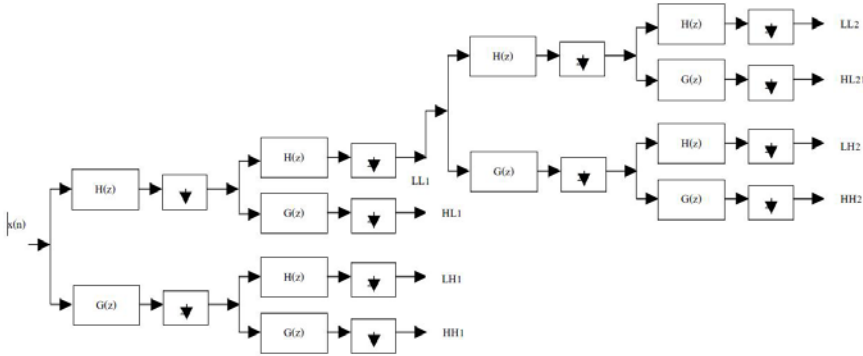


Fig. 2. Four channel perfect analysis filter bank

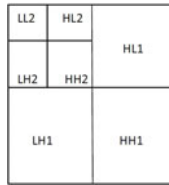


Fig. 3. Pyramidal Structure of Wavelet Analysis of Image

wavelet functions that scan details in horizontal $\psi(I)(x, y) = \psi(x)\psi(y)$, vertical $\psi(II)(x, y) = \psi(x)\psi(y)$, and diagonal directions: $\psi(III)(x, y) = \psi(x)\psi(y)$. This may be represented as a four-channel perfect reconstruction filter bank as shown in Fig. 2.

Now, each filter is 2D with the subscript indicating the type of filter (HPF or LPF) for separable horizontal and vertical components. The resulting four transform components consist of all possible combinations of high and low pass filtering in the two directions. By using these filters in one stage, an image can be decomposed into four different bands also known as sub-band coding. Reconstruction of the original image is accomplished by upsampling, filtering, and summing the individual subbands. There are three types of detail images for each resolution: horizontal (HL), vertical (LH), and diagonal(HH). The operations can be repeated on Low - Low band using the second stage of identical filter bank. Thus a typical 2-D DWT used in image compression [2] will generate the hierarchical pyramidal structure can be generate of the form as in Fig 3.

3 Multilayer Perceptron Model with Back Propagation learning for Image Compression

The classical model of MLP neural networks consists of three parts: an input layer, a hidden layer and an output layer. The principle of MLP neural networks for image compression is followed: the number of neurons in input layer and

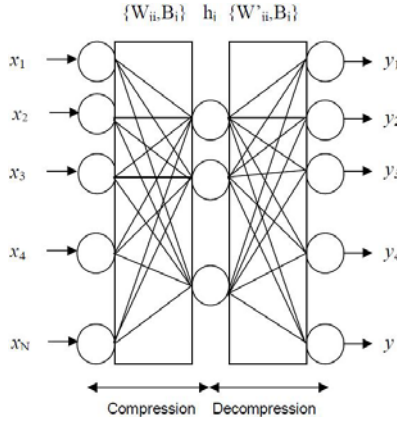


Fig. 4. Neural Network to use as Image Compression Decompression system

output layer is the same N , which is more than the number of neurons in hidden layer K . The neural networks are trained using the back-propagation learning algorithm: the image data is presented to the input of the network as column wise, in turn the weight and bias parameters of NN are modified until the mean-square difference between the output and the input is minimized. As a consequence, the weights and bias between input and hidden layer of the networks acts as a coder on the image data to be compressed, while the weights and bias between hidden and output layer is able to reconstruct the image, i.e. the decoder with the certain error. It must be observed that the trained neural network can be used without change on all kinds of images, thanks to its generalization property. The number of neurons in the hidden layer is significant for compression rate (CR)

$$\text{[9]: CR} = N/K$$

A $(K \times N)$ dimensional matrix W is the representation of the connection weights between neurons of input layer and hidden layer as is shown in the Fig.4, and bias of the j th neuron in hidden layer is denoted by $b_j : 1 \leq j \leq K$. Similarly, W represents the connection weights between neurons of hidden layer and output layer, and $\hat{b}_i : 1 \leq i \leq N$ is bias for the i^{th} neuron in hidden layer. The main task here is to find matrix W so that deviation between the input vector X and the reconstructed vector Y is minimized and expressed through mean square error (MSE). The neurons of the hidden layer are used as the components on which the input blocks will be projected. This operation can be formulated as 1.

$$h_j = f\left(\sum_{i=1}^N W_{ji}x_i + b_j\right) \quad 1 \leq j \leq K \quad (1)$$

The output of i -th neuron in output layer is:

$$y_i = \hat{f}\left(\sum_{j=1}^K \hat{W}_{ij}h_j + \hat{b}_{ij}\right) \quad 1 \leq i \leq N \quad (2)$$

In this paper, linear function is used as transfer function $f(\cdot)$ of hidden layer [7], and linear function as the transfer function $\hat{f}(\cdot)$ of output layer. To the original image with $n \times n$ pixels, it is necessary to divide the image into M non-overlapping blocks before compression, and each block contains $p \times p$ pixels.

The number of neurons in input layer and output layer is $p \times p$. Shown in Fig. 2, we regard the pixels of each block after pre-treatment as input data of the trained BP networks. Then the blocks are integrated to a complete image in the output terminal, consequently accomplishing the image reconstruction. According to the effect of image reconstruction using typical BP algorithm, there is a higher compression rate with fewer neurons in hidden layer; however, the quality of reconstructed image can be improved through, increasing the number of neurons in hidden layer.

4 Overview of Linear Programming

In canonical forms, the linear programming [1] the linear system of equations is described is in two forms

- i) The standard form
- ii) The slack form

In standard form, the linear function maximization is in from the linear inequalities and in slack form the linear function maximization is in from the linear equalities. The general form of expressing linear equation is:

Say for two variables x_1 and x_2 is

Maximize $x_1 + x_2$

subject to

$$a_1 * x_1 + a_2 * x_2 < b_1 \tag{3}$$

$$a_3 * x_1 + a_4 * x_2 > b_2 \tag{4}$$

...

...

$$a_5 * x_1 + a_6 * x_2 > b_2 \tag{5}$$

$$x_1, x_2 > 0 \tag{6}$$

We call any setting of variables x_1 and x_2 that satisfy all the constraints eq.3, 4, 5 is a feasible solution to linear program.

If we draw the graph of constraints (x_1, x_2) in Cartesian co-ordinate system as in fig 5

We obtained the set of feasible solutions and the convex region in 2-D space (shown enclosed area ϵ) i.e. feasible region. In this feasible region by evaluating

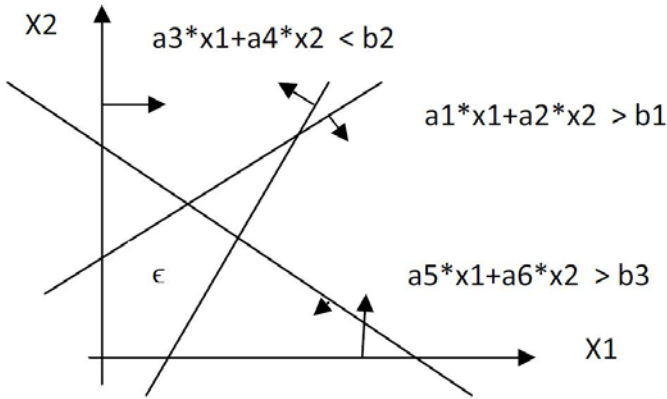


Fig. 5. Feasible region ϵ

the objective function i.e. here maximization of x_1+x_2 for all the available values in feasible region. The best possible solution in feasible region is known as an optimal solution and the solution of the function for which the result is desired (say) known as optimization function. One point should need to be understood here that the optimization function may be either maximum or minimum of the variables.

In general we cannot easily plot the graph for linear programs as the variables increases the complexity increases. Since feasible region has extremely large no. of solutions for the linear equations thus there is always need to devise the method for best possible solution i.e. the optimization of objective function. Linear programs are useful for modelling an approximate solution for the graph, combinatorial and network flow problems.

5 Proposed Method

For the standards like JPEG image compression schemes, the use of bigger quantization step sizes at high compression ratios, usually cause severe blockiness and other coding artifacts. At medium compression ratios, DCT-based image compression systems perform very well and its speed and compression capabilities are also good. But at higher compression ratios, image quality degrades as well as annoying artifacts results from the block-based DCT scheme. However, DWT based compression enables high compression ratio while maintaining good quality. In [10] Wavelet based coding schemes there is no need to blocked the image hence these schemes are more robust to transmission and coding errors and also they provide substantial improvement in picture quality at high compression ratios because of the overlapping basis functions and better energy compaction property of wavelet transforms.

A neuro-wavelet based hybrid model for image compression [3], combines the advantage of both the wavelet transforms and neural networks. Images are decomposed using wavelet filters into a set of sub bands with different resolution

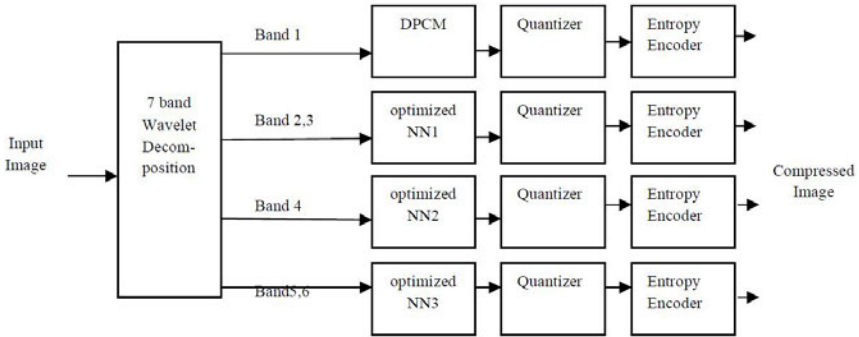


Fig. 6. Image Compression scheme based on Neuro Wavelet Hybrid Model

corresponding to different frequency bands. For various sub bands the different quantization and coding schemes are used based on their statistical properties. Initially the image is sub-band coded to two levels. The coefficients in lowest frequency band LL2 are compressed by differential pulse code modulation (DPCM) because they are extremely important for reconstruction so a lossless compression is required here and the coefficients in higher frequency bands are compressed using different neural networks. The advantage of having neural networks as Multi-Layer Perceptron with the hebbian learning rule brings the principle components of image.

The Proposed model is:

The image compression system using the wavelet transform and neural networks [6] is shown in Figure 6. The system first computes the seven-band wavelet decomposition of the input image. The lowest frequency band LL2 is most crucial band with regard to information point of view. To reconstruct the image from its data this band is very crucial. Hence LL2 i.e. Band1 is coded using DPCM i.e. Lossless. The remaining three frequency bands are coded using three different optimized neural networks. These optimized Neural Networks act as principal component extractor, here firstly every subband is divide into 8x8 block and hence each block is normalized. These three different NNs having configurations as

NN1 for Band 2,3 - 64 : 8 : 64

NN2 for Band 4 - 64 : 4 : 64

NN3 for Band 5,6 - 64 : 1 : 64

The hidden layer outputs of these NN are the principal components of the sub-image as well as the compressed data. The complete process for the NN1 is explained as:

These 8x1 compressed output values are fed as input to the decompressor

If we represent our Compression weight matrix as $W1$, Compression Bias Matrix as $B1$, Decompression weight matrix as $W2$ and Decompression Bias matrix as $B2$

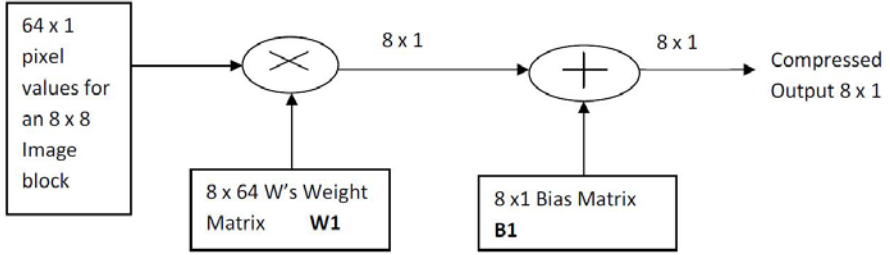


Fig. 7. Neural Network as Image Compressor

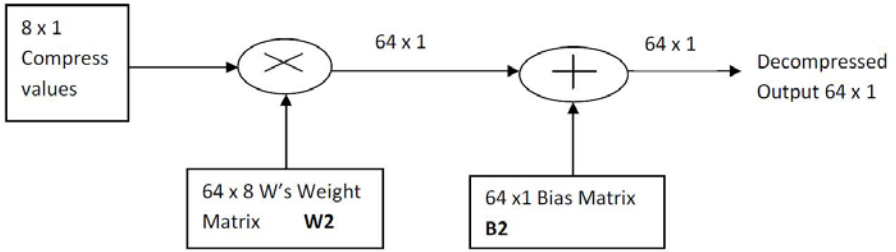


Fig. 8. Neural Network as Image Decompressor

Thus complete process as shown above, we can represent in the equation as

$$\text{Compression Stage} \quad W1 * X + B1 = X0 \quad (7)$$

$$\text{Decompression Stage} \quad W2 * X0 + B2 = X_{recon} \quad (8)$$

Where X is 64×1 input (i.e. 8×8 image block) and X_{recon} as reconstructed image at receiver, and $X0$ is 8×1 compressed output.

If we have trained our neural network for the exact Reconstruction of image i.e. $X_{recon} = X$ for every input block, then we can represent the complete process for no loss is

$$W2 * (W1 * X + B1) + B2 = X \quad (9)$$

But since we know this process can not be possible without loss thus we introduce error terms as slack variables and hence new final equation is:

$$W2 * (W1 * X + B1 + \epsilon1) + B2 + \epsilon2 = X \quad (10)$$

Where the $\epsilon1$ is error from input to hidden layer of MLP and $\epsilon2$ is error from hidden to output layer of MLP.

Hence we define here two terms first correction factor for weights as:

$$Cf_{wj} = \text{Min} \left(\sum_{i=1}^N \epsilon1_{ij} \right) = \text{Min} (\epsilon1_{1j} + \epsilon1_{2j} + \epsilon1_{3j} + \epsilon1_{4j} \dots + \epsilon1_{Nj}) \quad (11)$$

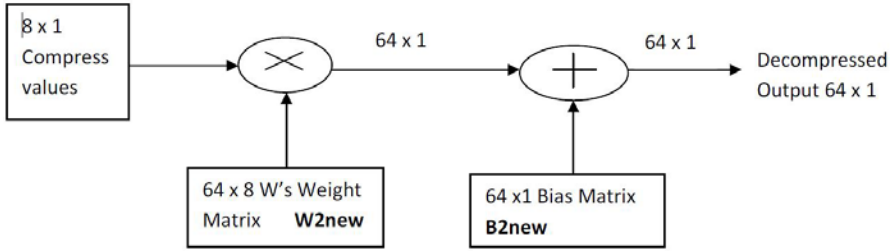


Fig. 9. New optimized model for NN as Image Decompressor

i.e. an objective function in linear programming terms where $j = 1, 2, 3 \dots N$ (The total no. of such 8x8 block).

And second correction factor for bias as

$$Cf_{bj} = \text{Min} \left(\sum_{i=1}^N \epsilon_{2ij} \right) = \text{Min} (\epsilon_{21j} + \epsilon_{22j} + \epsilon_{23j} + \epsilon_{24j} \dots + \epsilon_{2Nj}) \quad (12)$$

i.e. another objective function in linear programming terms

where $j = 1, 2, 3 \dots N$ (The total no. of such 8x8 block).

and the errors between the blocks is given as

$\epsilon_1 = \delta(X)$ i.e the some very small value of input X

$\epsilon_2 = \delta(X_{recon} - X)$ i.e the some very small value of error. ie. Output - Target

Hence the new Decompression Model is

Where

$$W2new_j = W2_j + Cf_{wj} \quad (13)$$

$$B2new_j = B2_j + Cf_{bj} \quad (14)$$

for all $j = 1, 2, 3 \dots N$

6 Computer Simulation Results

The experimental results of the proposed method are presented subjectively and objectively. Experiments were conducted using the images lena, pepper, house, goldhill, and woman of size 256 x 256, with 256 gray levels. Image was decomposed using Daubechies 4- coefficient filter (DAUB 4), 6-coefficient filter (DAUB 6), 18 coefficient filter (DAUB 18), 5 coefficients Coiflet filter (COIF5) and 9-coefficients Biorthogonal filter (BIOR9) for different comparisons. Band-1 is coded using DPCM, Band-2 and 3 is coded using a neural network with eight units in the input and the output and 6 hidden units i.e 8-6-8 neural network. Band-4 is coded using a 8-4-8 neural network, and band-5 and 6 using 16-1-16 network. The reason behind to having different lies in the fact that the components in band 2,3 are



Fig. 10. a) Original Image b) Reconstructed with Optimization c) Reconstructed without optimization

Table 1. NN Training with Lena image and compression for Lena

Wavelet Filter	bpp without Huffman encoding	bpp with Huffman encoding	MSE without optimization	MSE with optimization	PSNR without Optimization(dB)	PSNR with optimization (dB)
DAUB4	0.188	0.1380	97.29	34.68	28.25	32.73
DAUB6	0.188	0.1382	94.42	46.89	28.38	31.42
DAUB18	0.200	0.1384	83.38	60.96	28.92	30.28
COIF5	0.200	0.1382	73.46	61.38	29.47	30.25
BIOR9	0.198	0.1379	75.00	42.56	29.38	31.84

Table 2. NN Training with Lena image and compression for Peppers

Wavelet Filter	bpp without Huffman encoding	bpp with Huffman encoding	MSE without optimization	MSE with optimization	PSNR without Optimization(dB)	PSNR with optimization (dB)
DAUB4	0.188	0.1380	126.79	36.147	27.10	32.55
DAUB6	0.188	0.1382	120.56	46.034	27.33	31.50
DAUB18	0.200	0.1384	85.72	53.22	28.80	30.87
COIF5	0.200	0.1382	119.42	60.26	27.36	30.33
BIOR9	0.198	0.1379	125.33	48.314	27.15	31.29

more important than band 5,6 thus are represented with more bits. Scalar quantization and Huffman coding was then used on the coefficients of hidden layer to gain more compression. The metric used for comparison of original image f and decoded image \hat{f} is the PSNR (peak signal-to-noise ratio) value. In general, the larger PSNR value, the better is the reconstructed image qualitatively. The mathematical formulae for the computation of MSE & PSNR is

$$MSE = \sum_{m=1}^M \sum_{n=1}^N [pix_{org}(m, n) - pix_{comp}(m, n)]^2 \quad (15)$$

Table 3. Proposed algorithm comparison for different test images for NN training and testing with same test image alongwith standard approaches

Image	bpp	MSE	PSNR (dB)	Time(sec) for training the NN For MSE 0.02	PSNR (dB) with JPEG	PSNR (dB) with JPEG2000
Lena	0.1380	34.92	32.7	0.37 / 135 Epochs	30.654	30.512
Peppers	0.1392	48.20	31.3	0.41/142 Epochs	29.732	31.329
GoldHill	0.1396	49.33	31.2	0.41 / 142 Epochs	28.374	31.112
House	0.1385	31.84	33.1	0.38 / 152 Epochs	29.076	32.980
Woman	0.1392	15.24	36.3	0.36 / 135 Epochs	28.209	32.324

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (16)$$

7 Conclusion

This paper presented an optimized neuro-wavelet based image compression. This image compression method are further compared for the various kind of wavelets used. Comparison results show that DAUB18 performed better than other filter function. Representation of neural network as linear programming used to optimize the network greatly enhances the performance of the system.

References

1. Rivest, R.L., Stein, C., Leiserson, C.E., Cormen, T.H.: Introduction to algorithms, 2nd edn. MIT Press, Cambridge (2001)
2. Averbuch, A., Lazar, D., Israeli, M.: Image compression using wavelet transform and multiresolution decomposition. IEEE Trans. Image Processing 5, 4–15 (1996)
3. Cherkassky, V., Perhi, K., Denk, T.: Combining neural network and the wavelet transform for image compression. In: Proceeding of IEEE Intl. Conf., pp. 637–640 (1993)
4. Namphoi, A., Arozullah, M., Chin, S.: Higher order data compression with neural networks. In: IJCNN9, Seattle, vol. 9, pp. 55–60 (1999)
5. Kovacevic, J., Vetterli, M.: Wavelets and subband coding. Prentice Hall, Englewood Cliffs (1995)
6. Singh, V., Rajpal, N., Murthy, K.S.: A neuro-wavelet model using vector quantization for image compression. Computer Society of India 38(1), 10–19 (2008)

7. Kawato, M., Sonehars, N., Miyake, S.: Image compression using wavelet transform and multiresolution decomposition. In: Proc. IJCNN, Washington, DC, pp. 35–41 (1989)
8. Remelhart, D.E.: Learning internal representations by back propagating errors. *Nature* 1(1), 533–536 (1986)
9. Ramponi, G., Marsi, S., Sicuranza, G.L.: Improved neural structures for image compression. In: Proc IEEE Int. Conf. on Acoust, Speech and Signal Processing, pp. 281–284 (1991)
10. Premaraju, S., Mitra, S.: Efficient image coding using multi resolution wavelet transform and vector quantization. In: Image Analysis and Interpretation, pp. 135–140 (1996)
11. Vilovic, I.: An experience in image compression using neural networks. In: 48th International Symposium ELMAR, pp. 95–98 (2006)

Writer Based Handwritten Document Image Retrieval Using Contourlet Transform

M.S. Shirdhonkar¹ and Manesh B. Kokare²

¹ B. L. D. E. A's, College of Engineering, Bijapur, India

² S.G.G.S, Institute of Engineering and Technology, Nanded, India
ms_shirdhonkar@rediffmail.com,
mbkokare@sngs.ac.in

Abstract: Many techniques have been reported for handwritten based document image retrieval. This paper proposes a novel method by using Contourlet Transform (CT) for feature extraction of document images which achieves high retrieval rate. The handwriting of different people is often visually distinctive; we take a global approach based on texture analysis, where each writer's handwriting is regarded as a different texture. The Canberra distance is used as similarity measure in proposed system. A retrieval result with proposed method is very promising with precision and recall as compared to the existing system.

Keywords: Document image analysis and retrieval, CT, document similarity measurement, handwritten documents, handwriting classification and retrieval, writer identification.

1 Introduction

This article is focused on handwritten document analysis for experts. The method that we propose, allows the retrieval of digitized handwritten document images according to writers. Many organizations currently use and dependent on document image database. Searching for relevant document from large complex document image repositories is a central problem in document image analysis and retrieval. Document image retrieval provides user with an efficient way to identify a questioned document from among a large set of known database [1]. In this paper, we have proposed automatic handwritten document image retrieval based on writer which is independent of the textual contents. Document image retrieval is a very attractive field of research with the continuous growth of interest and increasing security requirements for the development of the modern society. Our objective is to find writer based document image retrieval. We use multi-channel spatial filtering techniques to extract texture features from handwriting images. There are many available filters in multi-channel techniques. In this paper the writer retrieval is based on the definition of features that can extract over an entire textual handwriting sample. Hence, we propose a global approach of writer based handwritten document image retrieval using Contourlet transform.

The need for searching scanned handwritten documents are involved in application such as collection dating, works genesis, critical edition, document authentication. In 2003, Srihari et al. [2], have realized the importance of handwritten document retrieval and have presented their retrieval system dedicated to forensics applications such as writer identification. In 2004, Schomaker and Bulacu [3], computes a code book of connected component contours from an independent training set and employs the probability density function of the unknown writing to identify its author. In 2005, Bensefia et al [4], uses local features based on graphemes that are produced by a segmentation algorithm based on the analysis of the minima of the upper contour. In 2006, Pareti and Vincent [5], models the distribution of patterns occurring in handwriting texts by zipt law, the respective zipt curve charactering the writer. In 2007, G.Joutel et al [6], proposed curvelets based queries for CBIR applications in handwriting collections, in this method curvelet coefficients are used as representation tool for handwriting when searching in large manuscripts databases by finding similar handwritten samples. In 2008, Siddiqi and Nicole [7], proposed effective method for writer identification in handwritten documents. This method is based on identify the writing style of an author and then extracting the forms that a writer used frequently as he draw the characters. In 2009, Siddiqi and Nicole [8], proposed a set of chain code based features for writer recognition. This method is based on finding the contours of a handwritten image and extracting a set of chain code based histograms at the global as well as local levels. In 2010, Liangshuo Ning et. al[9], proposed effective method for writer identification using multiscale gaussian markov random fields. Our work differs from those one because it uses a wavelet based tool, the Contourlet, which has never been used in this context.

The main contribution of this paper is that, we have proposed a novel writer based handwritten document image retrieval using Contourlet transform. Unique properties of CT like directionality and anisotropy made it a powerful tool for feature extraction of images in the database. Handwritten document matching is performed using Canberra distance. The experimental results of proposed method were satisfactory and give better results as compared with earlier approaches. The rest of paper is organized as follows. Section 2, discusses the proposed system. Handwritten document image retrieval phase is presented in section 3. In section 4 the experimental results are presented and finally section 5, conclude the work.

2 Proposed System

The objective of the proposed work is to study the use of edge and texture orientations as image features in document image retrieval. The basic architecture of the system is shown in Fig. 1. An improved method based on Contourlet transform is proposed in this work. There are two issues in building the proposed system

- Every image in the image data base is to be represented efficiently by extracting significant features.
- Relevant images are to be retrieved using similarity measures between query and every image in the image database

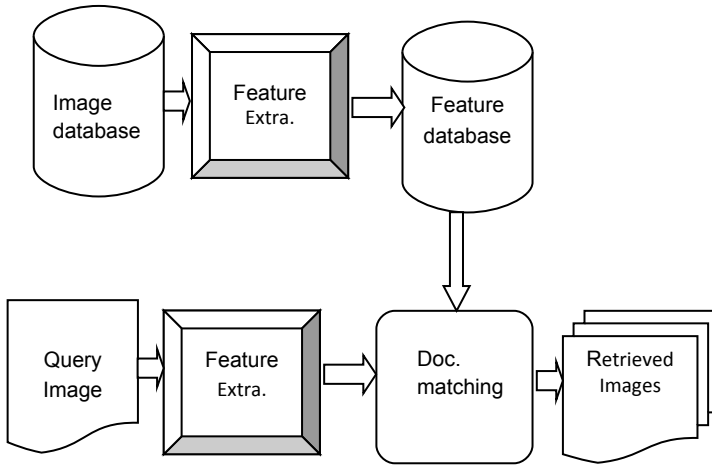


Fig. 1. System Architecture for Proposed System

The performance of the proposed system can be tested by retrieving the desired number of the handwritten document images from the document image database. The average retrieval rate is the main performance measures in the proposed system.

2.1 Feature Extraction

Features used for the document image retrieval task are mainly global features which are based on statistical measurements, extracted from the whole block of text to be identified. These features can be broadly classified into two families:

- Features extracted from texture: the document image is simply seen in this case as an image and not as handwriting. For example, application of the multi-channel Gabor filtering technique, co-occurrence matrices, Curvelet transform and Contourlet transform were considered.
- Structural features: in this case the extracted features attempt to describe some structural properties of the handwriting. For example the average height, the average width, and the average slope.

In our proposed approach uses Contourlet transform technique is implemented to obtain texture features. Multiscale and time-frequency localization of an image is offered by wavelets. But, wavelets are not effective in representing the images with smooth contours in different directions. Contourlet Transform (CT) addresses this problem by providing two additional properties viz., directionality and anisotropy [10], which are defined as:

- Directionality: The representation should contain basis elements oriented at a variety of directions, much more than the few directions that are offered by wavelets.
- Anisotropy: To capture smooth contours in images, the representation should contain basis elements using a variety of elongated shapes with different aspect ratios.

Contourlet transform is a multiscale and directional representation that uses first a wavelet like structure for edge detection and then a local directional transform for contour segment detection. In the double filter bank structure, Laplacian Pyramid (LP) is used to capture the point discontinuities and then followed by a directional filter bank (DFB), which is used to link these point discontinuities into linear structures. The Contourlets have elongated supports at various scales, directions and aspect ratios. This allows Contourlets to efficiently approximate a smooth contour at multiple resolutions [11]. In the frequency domain, the Contourlet transform provides a multiscale and directional decomposition. Contourlet transform is simple and flexible but it introduces redundancy (up to 33%) due to the LP stage. These properties of CT i.e. directionality and anisotropy made it a powerful tool for content based image retrieval (CBIR).

2.1.1 Laplacian Pyramid Decomposition

To obtain multiscale decomposition LP is used. LP decomposition at each level generates a down sampled low pass version of the original image and difference between the original and the prediction that results in a band pass image. The LP decomposition is shown in Fig. 2. In LP decomposition process, H and G are one dimensional low pass analysis and synthesis filters respectively. M is the sampling matrix. Here, the band pass image obtained in LP decomposition is then processed by the DFB stage. LP with orthogonal filters provides a tight frame with frame bounds equal to 1.

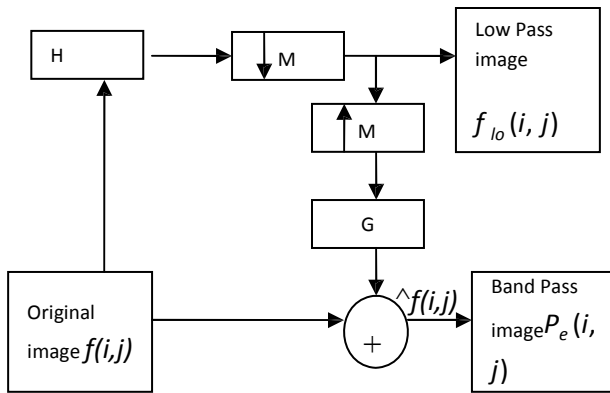


Fig. 2. LP Decomposition (One Level)

In LP decomposition of an image, $f(i,j)$ represent the original image and its low pass filtered version is $f_{lo}(i,j)$ and the prediction is $\hat{f}(i,j)$. The prediction error is given by

$$P_e(i,j) = f(i,j) - \hat{f}(i,j) \tag{1}$$

The directional decomposition is performed on $P_e(i, j)$ as it is largely decorrelated and requires less number of bits than $f(i, j)$. In Equ. 1, $P_e(i, j)$ represents a band pass image. Further decomposition can be carried by applying Eq. 1 on $f_{l0}(i, j)$ iteratively to get $f_{l1}(i, j), f_{l2}(i, j) \dots f_{ln}(i, j)$, where ‘n’ represents the number of pyramidal levels. In LP reconstruction the image is obtained by simply adding back the difference to the prediction from the coarse image.

2.1.2 DFB Decomposition

DFB is designed to capture the high frequency content like smooth contours and directional edges. Several implementations of these DFBs are available in the literature. This DFB is implemented by using a k-level binary tree decomposition that leads to 2^k directional sub-bands with wedge shaped frequency partitioning as shown in Fig. 3. But the DFB used in this work is a simplified DFB [13], which is constructed from two building blocks. The first one is a two-channel quincunx filter bank with fan filters. It divides a 2-D spectrum into two directions, horizontal and vertical. The second one is a shearing operator, which amounts to the reordering of image pixels. Due to these two operations directional information is preserved. This is the desirable characteristic in system to improve retrieval efficiency. Band pass images from the LP are fed to DFB so that directional information can be captured. The scheme can be iterated on the coarse image. This combination of LP and DFB stages result in a double iterated filter bank structure known as Contourlet filter bank, which decomposes the given image into directional sub-bands at multiple scales

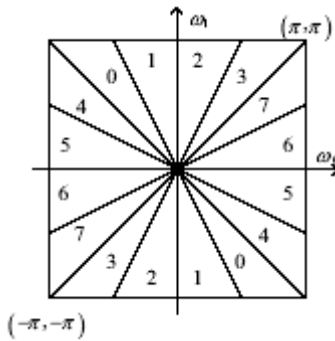


Fig. 3. DFB Frequency Partitioning

3 Handwritten Document Image Retrieval

3.1 Feature Database Creation

To conduct the experiments, each image from database is decomposed using CT with a 4 level (0, 2, 3, 4) LP decomposition. At each level, the numbers of directional subbands are 3, 4, 8 and 16 respectively. These parameters results in a 62 dimensional feature vector. To construct the feature vectors of each image in the database, the

energy and standard deviation were computed separately on each subband and the feature vector was formed using these two parameter values and normalized [12]. The retrieval performance with combination of these two feature parameters always outperformed that using these features individually. The energy (E_k) and (σ_k) standard deviation of k^{th} subband is computed as follows .

$$E_k = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N |W_k(i, j)| \tag{2}$$

$$\sigma_k = \left[\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (W_k(i, j) - \mu_k)^2 \right]^{\frac{1}{2}} \tag{3}$$

Where $W_k(i, j)$ is the k^{th} contourlet decomposed subband, $M \times N$ is the size of contourlet decomposed subband, and μ_k is the mean of the k^{th} subband. The resulting feature vector using energy and standard deviation are $\bar{f}_E = [E_1 \ E_2 \ \dots \ E_n]$ and $\bar{f}_\sigma = [\sigma_1 \ \sigma_2 \ \dots \ \sigma_n]$ respectively. So combined feature vector is

$$\bar{f}_{\sigma E} = [\sigma_1 \ \sigma_2 \ \dots \ \sigma_n \ E_1 \ E_2 \ \dots \ E_n] \tag{4}$$

We normalized the feature vector \bar{f}_E and \bar{f}_σ by applying the following statistical normalization method given in Equ. (5) and Equ. (6) respectively.

$$\bar{f}_{VE} = \frac{\bar{f}_E - \mu_{\bar{f}_E}}{\sigma_{\bar{f}_E}} \tag{5}$$

$$\bar{f}_{V\sigma} = \frac{\bar{f}_\sigma - \mu_{\bar{f}_\sigma}}{\sigma_{\bar{f}_\sigma}} \tag{6}$$

Where $\mu_{\bar{f}_E}$, $\mu_{\bar{f}_\sigma}$ and $\sigma_{\bar{f}_E}$, $\sigma_{\bar{f}_\sigma}$ are the mean and the standard deviation of \bar{f}_E , \bar{f}_σ respectively. Finally, the resultant feature vector will be the combined normalized vector f_V .

$$f_V = [\bar{f}_{V\sigma}, \bar{f}_{VE}] \tag{7}$$

For the creation of feature database, the above procedure is repeated for all the images in the database and these feature vectors are stored in the feature database

3.2 Document Image Matching

There are several ways to work out the distance between two points in multidimensional space. We have used Canberra distance as similarity measure. If x and y are the feature vectors of the database and query image respectively, and have dimension d , then the Canberra distance is given by Equ. 8

$$Canb(x, y) = \sum_{i=1}^d \frac{|x_i - y_i|}{|x_i| + |y_i|} \tag{8}$$

The average retrieval rate for the query image is measured by counting the number of handwritten document images from the same category which are found in the top ‘N’ matches.

4 Experimental Results

4.1 Image Database

The handwritten document images were collected using either black ink or blue ink (No pen brands were taken into consideration), on a white A4 sheet of paper, with one handwriting document image per page. A scanner subsequently digitized the document, contained on each page, with a resolution of 256 x 256 pixels. A number of experiments were carried out to show the effectiveness of the proposed system. A group of 15 writers are selected for 16 specimen handwritten document which make the total of 15x16=240 handwritten document database. Examples of handwriting document by these writers are shown in figure 4, for the purpose of retrieval.

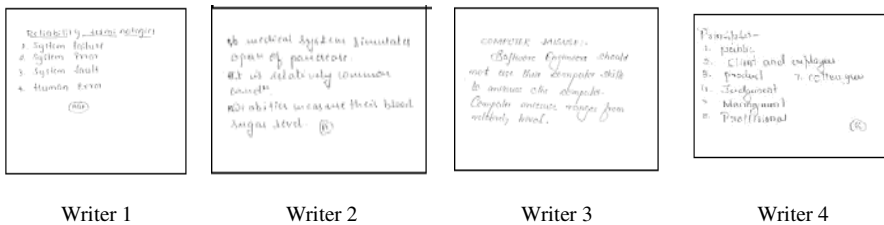


Fig. 4. Examples of handwriting documents of four different writers

4.2 Retrieval Performance

For each experiment, one image was selected at random as the query image from each writer and thus retrieved images were obtained. For performance evaluation of the document image retrieval system, it is significant to define a suitable metric. Two metrics are employed in our experiments as follows.

$$Recall = \frac{\text{Number of relevant documents retrieved}}{\text{Number of relevant documents in database}} \tag{9}$$

$$Precision = \frac{\text{Number of relevant documents retrieved}}{\text{Number of documents retrieved}} \tag{10}$$

Results correspond to precision and recall rate for a Top1, Top 2, Top 5, and Top 10, and Top 15. The comparative retrieval performance of the proposed system on the database using CT features is shown in Table 1.

Table 1. Average Retrieval Performance

No. of Top matches	Curvelet Transform		Contourlet Transform	
	Precision %	Recall %	Precision %	Recall %
Top 1	100	6.25	100	6.25
Top 2	100	12.5	95	11.87
Top 5	88	27.5	94	29.35
Top 8	69	34.37	76	38.12
Top 10	63	36.85	72	45.00
Top 15	49	46.25	61	56.87

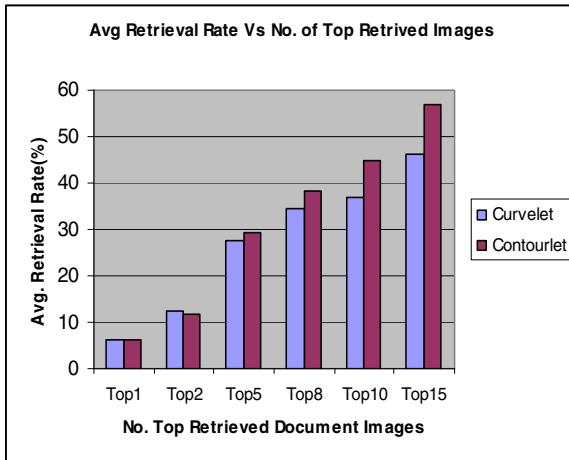


Fig. 5. Comparative average retrieval rate using Curvelet and Contourlet transform

Retrieval performance of the proposed method is compared using Curvelet transform technique[6]. We evaluated the performance in terms of average rate of retrieving images as function of the number of top retrieved images. Figure 5, shows graph illustrating this comparison between Curvelet transform and Contourlet

transform according to the number of top matches considered for database. From Fig. 5, it is clear that the new method is superior to Curvelet transform. The proposed method always outperforms the earlier method. The comparative performances in terms of average retrieval rate are shown in Fig. 5. To retrieve images from the database those have a similar writer to the original request. In Fig. 6, retrieval example results are presented in a list of images having a query image.

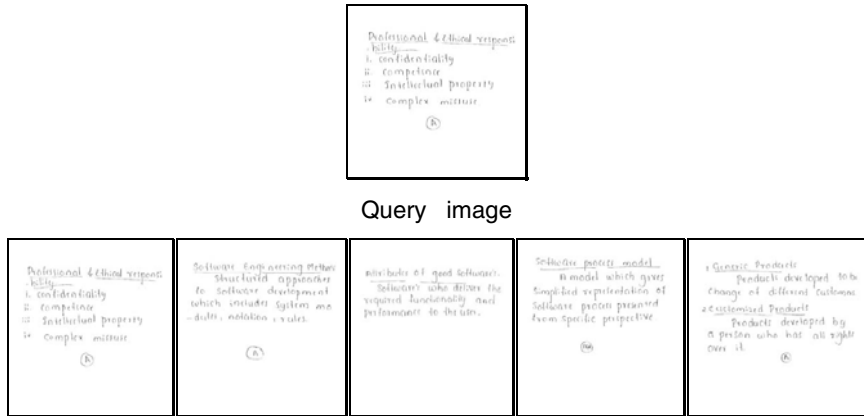


Fig. 6. List of the five most similar retrieved handwritten document images from the database

5 Conclusions

We have described a new approach towards writer based handwritten document image retrieval using Contourlet transform. This approach is based on the observations of the handwriting of different writer is visually distinctive and global approach based on texture analysis has been adopted. Features were extracted from handwriting images using Contourlet transform technique. Handwritten document matching was performed using the Canberra distance. Retrieval results with proposed method are very promising with precisions and recalls as compared to the existing system.

Acknowledgments. The authors would like to appreciate all writers who gave permission to use their handwritten documents in this study.

References

1. Shirdhonkar, M.S., Kokare, M.: Handwritten Document Image Retrieval. In: Proceeding of 3rd International Conference on Computer Modeling and Simulation (ICCMS 2011), Mumbai, India, pp. VI-506–VI-510 (2011)
2. Zhang, B., Srihari, S.N.: Binary Vector Dissimilarly Measures for Handwriting Identification. In: Document Recognition and Retrieval. SPIE, pp. 28–38 (2003)

3. Schomaker, L., Bulacu, M.: Automatic Writer Identification Using Connected Component Contours and Edge Based Features of Uppercase Western Script. *IEEE Trans. of Pattern Analysis and Machine Intelligence* 26(6), 787–798 (2004)
4. Bensefia, A., Paquet, T., Heutte, L.: A Writer Identification and Verification System. *Pattern Recognition Letters* 26(13), 2080–2092 (2005)
5. Pareti, R., Vincent, N.: Global Method Based on Pattern Occurrences For Writer Identification. In: *Proc. of the 10th International Workshop on Frontiers in Handwriting Recognition*, La Baule, France (2006)
6. Joutel, G., Eglin, V., Bres, S., Emptoz, H.: Curvelets Based Features Extraction of Handwritten Shapes for Ancient Manuscript Classification. In: *Proc. Document Recognition and Retrieval XIV, SPIE –IS&T Electronic imaging*. SPIE, vol. 6500, pp. 65000D-1–65000D-10 (2007)
7. Siddiqi, I., Vincent, N.: Combining Global and Local Features for Writer Identification in Handwritten Documents. In: *Proc. of the 11th International Conference on Frontiers in Handwriting Recognition*, Canada (2008)
8. Siddiqi, I., Vincent, N.: A Set of Chain Code Based Features for Writer Recognition. In: *10th International Conference on Document Analysis and Recognition*, pp. 981–985 (2009)
9. Ning, L., Zhou, L., You, X., Du, L., He, Z.: Multiscale Gaussian Markov Random fields for writer identification. In: *Proc. of International Conference on Wavelet Analysis and Pattern Recognition*, pp. 170–175 (2010)
10. Srinivasa Rao, C., Srinivas Kumar, S., Chatterji, B.N.: Content Based Image Retrieval Using Contourlet Transform. *ICGST-GVIP Journal* 7(3) (2007)
11. Po, D.D.Y., Do, M.N.: Directional Multiscale Modeling of Image Using the Countourlet Transform. *IEEE Transactions on Image Processing* 15(6), 1610–1620 (2006)
12. Kokare, M., Biswas, P.K., Chatterji, B.N.: Texture Image Retrieval Using New Rotated Complex Wavelets Filters. *IEEE Trans. on Systems, Man, and Cybernetics-Part B: Cybernetics* 35(6), 1168–1178 (2005)
13. Bamberger, R.H., Smith, M.J.T.: A Filter Bank for the Directional Decomposition of Images: Theory and design. *IEEE Transactions on Signal Processing* 40(4), 882–893 (1992)

Semantic Learning in Interactive Image Retrieval

Pushpa B. Patil¹ and Manesh Kokare²

¹ Dept. of Computer Science and Engg.

BLDEA's, V.P.Dr. P. G. Halakatti CET, Bijapur, India

² Dept. of Electronic and Telecommunication

S.G.G.S, Institute of Engineering and Technology, Nanded, India

pushpa_ms@rediffmail.com,

mbkokare@sngs.ac.in

Abstract. This paper presents content-based image retrieval frameworks with relevance feedback based on AdaBoost learning method. As we know relevance feedback (RF) is online process, so we have optimized the learning process by considering the most positive image selection on each feedback iteration. To learn the system we have used AdaBoost. The main significances of our system are to address the small training sample and to reduce retrieval time. Experiments are conducted on 1856 texture images to demonstrate the effectiveness of the proposed framework. These experiments employed large image databases and combined RCWFs and DT-CWT texture descriptors to represent content of the images.

Keywords: Content-Based Image Retrieval (CBIR), Relevance Feedback (RF), Rotated Complex Wavelet Filters (RCWFs), Dual Tree Complex Wavelet (DT-CWT), and Image retrieval.

1 Introduction

Recently, there is a huge collection of digital images are available. Since there is a development of the internet and availability of digital devices such as scanners, digital cameras etc. With these huge collections, it is very essential to have efficient and effective tools for searching, browsing and retrieval of images. To do these tasks CBIR was introduced in the early 1980. CBIR uses the low level features like color, texture and shape to retrieve the most similar images stored in the database. With these low level features, user's perception on the images can not be achieved. Since different users perception is different on same images. It is the big disadvantage of the CBIR. To overcome this, relevance feedback (RF) was introduced into CBIR in 1998[5]. There is good review on CBIR [1-4].

RF is an online process, which tries to learn the user perception interactively; initially RF was designed for text-based information retrieval systems. Later it was introduced into CBIR during mid 1990's, with the involvement of the user in the retrieval loop to reduce the "semantic gap" between query representation (low level features) and user perception (high level concepts). RF has been proved to provide effective and efficient retrieval performance improvement in CBIR systems through interactive learning based on the user feedback [5-6].

1.1 Related Work

Recently, many researchers began to consider the RF as a classification or semantic learning problem. That is a user provides positive and/or negative examples, and the systems learn from such examples to separate all data into relevant and irrelevant groups. Hence many classical machine learning schemes may be applied to the RF, which include decision tree learning [7], Bayesian learning [8]-[9], support vector machines [10], boosting [11] and so on. There is good review on RF in [12]. The process of learning is very difficult task in RF [12]-[14], due to the following reasons

- 1) Training data is very small, which is less than the dimension of the feature space. This makes difficult to apply most of the learning methods such as linear discriminate fisher classifier and relevance vector machine (RVM). Though the RVMs are sparser than the SVMs and use less number of kernel functions.

- 2) Training data is asymmetrical, which creates a too much imbalance between the relevant and irrelevant images.

- 3) In RF, for every iteration we have to perform both training and testing online, which takes more real time.

For visual representation of the images, we employed the global texture features presented in [18], which provides very efficient performance. Much of the work on RF uses the low-level representation using discrete wavelet transform (DWT) [15], Gabor filters [16] and co-occurrence matrix [19][20] for textures. In order to retrieve the general purpose images like artificial objects and natural scenes most of the time textural features are combined with color and shape to get better retrieval performance. However, they still suffers from the poor directional sensitivity, shift variant and, redundancy. From these combined features we may get better retrieval performance but not efficient one because as we increase number of features it increases the dimensionality of feature space. With such high dimensional feature space, RF may become impractical for even medium sized databases [14]. In order to store and process these high dimensional feature vectors it requires more memory space and time. So, to make our system efficient, we have to consider two factors namely time complexity and space complexity together with better retrieval performance. To overcome above problem, we propose to use the new rotated complex wavelet filters which gives both better and efficient retrieval performance.

1.2 Our Approach

In this paper we have used our earlier recent work [18] to extract more compact effective low level features, to improve the retrieval performance in terms of speed, storage and accuracy by using the rotated complex wavelet filters and dual tree complex wavelet transform jointly. To reduce the significant gap between low level feature and high level concepts, we have proposed a new RF approach and it is tested using AdaBoost. We found that proposed RF framework provides efficient retrieval performance in very few feedback iterations. A new relevance feedback approach, which is based on a ADABOOST uses the relevant and irrelevant examples. Our extensive experiments using proposed RF with AdaBoost on standard texture database show significant improvements with respect to retrieval performance.

The rest of the paper is organized as follows, we briefly discuss the dual-tree complex wavelet, and dual tree rotated complex wavelet in section 2. In section 3, we have, explained the concept of AdaBoost. In section 4, experimental results are discussed. Finally, the conclusion is given in section 5.

2 Image Descriptors

2.1 DT-CWT

Real DWT has poor directional selectivity and it lacks shift invariance. Drawbacks of the DWT are overcome by the complex wavelet transform (CWT) by introducing limited redundancy into the transform. But still it suffer from problem like no perfect reconstruction is possible using CWT decomposition beyond level 1, when input to each level becomes complex. To overcome this, Kingsbury [21] proposed a new transform, which provides perfect reconstruction along with providing the other advantages of complex wavelet, which is DT-CWT. The DT-CWT uses a dual tree of real part of wavelet transform instead using complex coefficients. This introduces a limited amount of redundancy and provides perfect reconstruction along with providing the other advantages of complex wavelets. The DT-CWT is implemented using separable transforms and by combining subband signals appropriately. Even though it is non-separable yet it inherits the computational efficiency of separable transforms. Specifically, the 1-D DT-CWT is implemented using two filter banks in parallel, operating on the same data. For d -dimensional input, a L scale DT-CWT outputs an array of real scaling coefficients corresponding to the lowpass subbands in each dimension. The total redundancy of the transform is 2^d and independent of L . The mechanism of the DT-CWT is not covered here. See [22] and [23] for a comprehensive explanation of the transform and details of filter design for the trees. A complex valued $\psi(t)$ can be obtained as

$$\psi(x) = \psi_h(x) + j \psi_g(x) \quad (1)$$

Where $\psi_h(x)$ and $\psi_g(x)$ are both real-valued wavelets. The impulse response of six wavelets associated with 2-D complex wavelet transform are illustrated in Fig. 1.

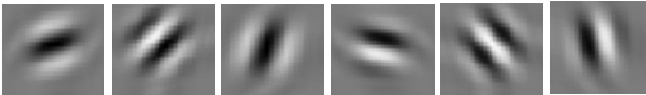


Fig. 1. Impulse response of six wavelet filters of complex wavelet.

2.1 DT-RCWF

Recently we have designed 2D- rotated complex wavelet transform [18]. Directional 2D RCWF are obtained by rotating the directional 2D DT-CWT filters by 45° so that decomposition is performed along new direction, which are 45° apart from

decomposition of CWT. The size of a newly obtained filter is $(2N-1) \times (2N-1)$, where N is the length of the 1-D filter. The decomposition of input image with 2-D RCWF followed by 2-D downsampling operation is performed up to the desired level. The computational complexity associated with RCWF decomposition is the same as that of standard 2-D DT-CWT, if both are implemented in the frequency domain. The set of RCWFs retains the orthogonality property. The six subbands of 2D DT-RCWF gives information strongly oriented at $(30^\circ, 0^\circ, -30^\circ, 60^\circ, 90^\circ, 120^\circ)$. The mechanism of the DT-RCWF is explained in our earlier work [18]. The 2D DT-CWT and RCWF provide us with more directional selectivity in the direction $\left\{ \begin{matrix} (+15^\circ, +45^\circ, +75^\circ, -15^\circ, -45^\circ, -75^\circ) \\ (0^\circ, +30^\circ, +60^\circ, +90^\circ, 120^\circ, -30^\circ) \end{matrix} \right\}$ than the DWT whose directional sensitivity is in only four directions $\{0^\circ, \pm 45^\circ, 90^\circ\}$. The impulse response of six wavelets associated with rotated complex wavelet transform is illustrated in Fig. 2.

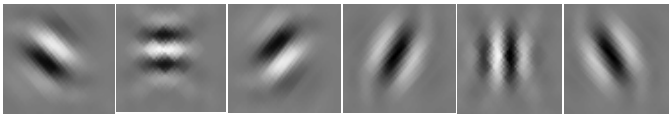


Fig. 2. Impulse response of six rotated complex wavelet filter

3 AdaBoost

AdaBoost was introduced in 1995 by Freund and Schapire [24] as an efficient algorithm of the ensemble learning field. It is used to boost the classification performances of a weak learner. It does this by combining a collection of weak classification functions to form a stronger classifier. AdaBoost combines iteratively the weak classifiers by taking into account a weight distribution on the training samples such that more weight is attributed to samples misclassified by the previous iterations.

Consider a two classification problem, in which the training data comprises input vectors X_1, \dots, X_N along with corresponding binary target variables t_1, \dots, t_n , where $t_n \in \{-1, 1\}$. each data point is given an associated weighting parameter w_n , which is initially set $1/N$ for all data points. We shall suppose that we have a procedure available for training a base classifier using weighted data to give a function $y(x) \in \{-1, 1\}$. At each stage of the algorithm, AdaBoost trains a new classifier using a data set in which the weighting coefficients are adjusted according to the misclassified data points. Finally, when the desired number of base classifiers has been trained, they are combined to form a committee using coefficients that give different weight to different base classifiers.

3.1 Image Retrieval

3.1.1 Feature Database Creation

To conduct the experiments, each image from database is decomposed using DT-CWT and DT-RCWF up to third level and two different sets of features were computed as follows.

To construct the feature vectors of each image in the database, we decomposed each image using DT-CWT and DT-RCWF up to third level. The Energy and Standard Deviation (STD) were computed separately on each subband and the feature vector was formed using these two parameter values. The retrieval performance with combination of these two feature parameters always outperformed that using these features individually [18]. The Energy (E_k) and (σ_k) Standard Deviation of k^{th} subband is computed as follows

$$E_k = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N |W_k(i, j)| \quad (5)$$

$$\sigma_k = \left[\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (W_k(i, j) - \mu_k)^2 \right]^{\frac{1}{2}} \quad (6)$$

Where $W_k(i, j)$ is the k^{th} wavelet-decomposed subband, $M \times N$ is the size of wavelet decomposed subband, and μ_k is the mean of the k^{th} subband. The resulting feature vector using energy and standard deviation are $\bar{f}_E = [E_1 \ E_2 \ \dots \ E_n]$ and $\bar{f}_\sigma = [\sigma_1 \ \sigma_2 \ \dots \ \sigma_n]$ respectively. So combined feature vector is

$$\bar{f}_{\sigma E} = [\sigma_1 \ \sigma_2 \ \dots \ \sigma_n \ E_1 \ E_2 \ \dots \ E_n] \quad (7)$$

3.2 Image Matching

We have randomly selected any one of the 1856 images as a query image from texture images. Query image is further processed to compute the feature vector as given in section 3.1. Canberra distance metric is used as a similarity measure. If x and y are the feature vectors of the database and query image, respectively, and have dimension d , then the Canberra distance is given by

$$\text{Canb}(x, y) = \sum_{i=1}^d \frac{|x_i - y_i|}{|x_i| + |y_i|} \quad (8)$$

4 Experiments

To evaluate the performance of a proposed system, we have used the Brodatz texture photographic album. The experiments were conducted using MATLAB 7.0 with Intel core2Duo, 1 GB RAM machine.

4.1 Image Database

The texture database used in our experiment consists of 116 different textures [18]. We used 108 textures from Brodatz texture photographic album, seven textures from USC database and one artificial texture. Size of each texture image is 512×512 . Each 512×512 image is divided into sixteen 128×128 non overlapping sub images, thus creating a database of 1856 texture images.

4.2 Performance Measures

For experimental results, it is significant to define a suitable metric for performance evaluation. We have used Average accuracy and it is defined as the percentage of relevant images of retrieved images among all relevant images in the database.

Experimental results are evaluated on 116 queries randomly selected from the texture database. The reported results of average accuracy are obtained by taking an average over the 116 queries texture database.

For each experiment, one image was selected at random as the query image from each category and thus the retrieved images were obtained. Then, the users were asked to identify those images that are related to their expectations from the retrieved images. These selected images were used as feedback images for next iteration. Finally, we have computed the average accuracy of all the categories in the database. Each image category contains 16 images. The feedback processes were performed 4 times.

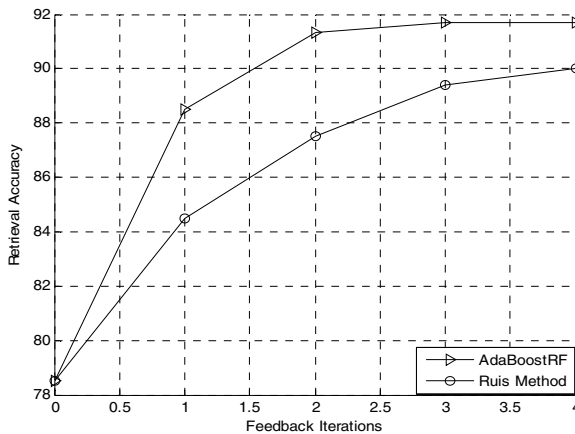


Fig. 3. Average accuracy versus iteration curves for texture images

Fig.3 describes detailed comparison of the average retrieval accuracy obtained using Rui[5] and AdaBoost on every feedback iteration of the randomly selected image from each category of texture database. The main observation of proposed RF using AdaBoost gives better retrieval performance comparing with Rui. method. From the Fig. 3, we observed that, there is a rapid increase in retrieval performance with each feedback iteration of proposed RF using AdaBoost learning algorithm. Retrieval performance is improved from 78.5% to 91.70% from first iteration to the fourth iteration using AdaBoostRF

We illustrated these observations using graph in Fig. 3 and results are also tabulated in table 1.

Table 1. Average accuracy on each feedback iterations for Texture database

Approach	CBIR	1 st iteration	2 nd iteration	3 rd iteration	4 th iteration
AdaBoostRF	78.5	88.52	91.32	91.70	91.70
Rui method	78.5	84.50	87.52	89.42	90.01

4.3 Image Retrieval Examples

We use an example to illustrate the performance improvement of the proposed approach in Fig. 4(a)-4(c) for texture database. Fig.4(a) is the result of CBIR using combined features (RCWT+DT-CWT), in which among top 20 images, 8 images belongs to the desired category (i.e images 1-6 and image 16, 20) and remaining 12 belongs to irrelevant category. So we got 50.0% retrieval precision from CBIR. Fig.4 (b)-(c) shows performance improvement of the proposed approach using the AdaBoost for texture database. From Fig. 4(b) to 4(c), we can observe that retrieval accuracy increasing from 81.25% to 93.75% from first iteration to second iteration of relevance feedback and it remains same in further iterations

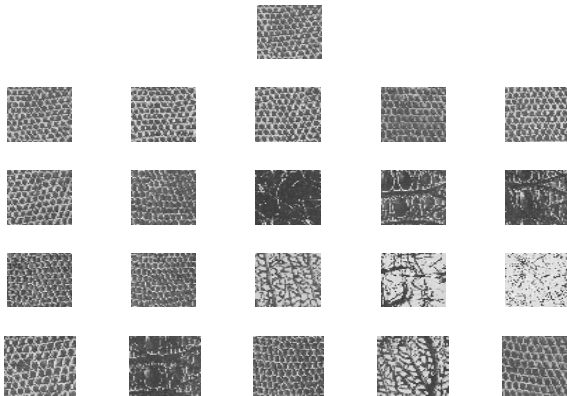


Fig. 4. (a) Result of CBIR using Combined features (RCWF +DT_CWT) (8/16)

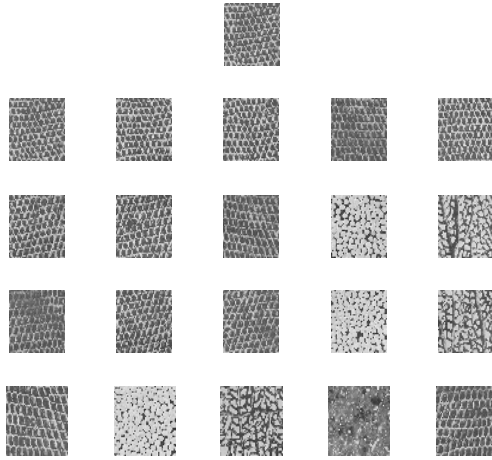


Fig. 4. (b) Result after first feedback iteration using AdaBoostRF(13/16)

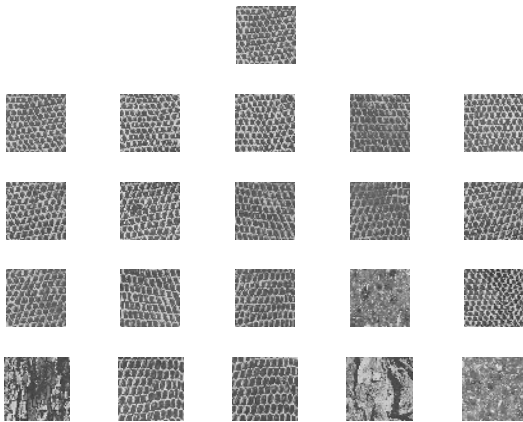


Fig. 4. (c) Result after second feedback iteration using AdaBoostRF (15/16)

5 Conclusion

In this paper, an active relevance feedback framework has been proposed to handle the small training data problem in RF and optimizing the testing set in order to reduce the retrieval time. The proposed relevance feedback framework is tested using AdaBoost with texture features. The RF framework is tested on large scale standard texture image database. The framework has demonstrated very promising retrieval accuracy. From experimental results, we found that RF using AdaBoost with combined texture features RCWF and DT-CWT gives better retrieval performance.

References

- [1] Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image Retrieval: Ideas, Influences, and Trends of the New Age. *ACM Computing Surveys* 40(2), article 5, 5:1–5:60 (2008)
- [2] Rui, Y., Hung, T.S., Chang, S.F.: Image retrieval: Current Techniques, Promising Directions and Open Issues. *J. Visual Comm. and Image Representation* 10, 39–62 (1999)
- [3] Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R.: Content –based Image Retrieval at the End of the Early Years. *IEEE Trans. Pattern Anal. Machine Intell.* 22(12), 1349–1380 (2000)
- [4] Kokare, M., Chatterji, B.N., Biswas, P.K.: A Survey on Current Content-based Image Retrieval Methods. *IETE J. Res.* 48(3&4), 261–271 (2002)
- [5] Rui, Y., Huang, T., Ortega, M., Mehrotra, S.: Relevance Feedback: A Power Tool in Interactive Content-Based Image Retrieval. *IEEE Transactions on Circuits and Systems for Video Technology* 8(5), 644–655 (1998)
- [6] Rui, Y., Huang, T.S., Mehrotra, S.: Content-based Image Retrieval with Relevance Feedback in MARS. In: *Proc. IEEE Int. Conf. on Image Proc.* (1997)
- [7] MacArthur, S.D., Brodley, C.E., Shyu, C.R.: Relevance Feedback Decision Trees in Content-based Image Retrieval. In: *Proc. IEEE Workshop Content-based Access of Image and Video Libraries*, pp. 68–72 (2000)
- [8] Cox, I.J., Miller, M.L., Minka, T.P., Papathomas, T.V., Yianilos, P.N.: The Bayesian Image Retrieval System, PicHunter: Theory, Implementation and Psychophysical Experiments. *IEEE Tran. on Image Processing* 9(1), 20–37 (2000)
- [9] Su, Z., Zhang, H., Li, S., Ma, S.: Relevance Feedback in Content-based Image Retrieval: Bayesian framework, Feature Subspaces, and Progressive Learning. *IEEE Trans. Image Process.* 12(8), 924–936 (2003)
- [10] Tong, S., Chang, E.: Support Vector Machines Active Learning for Image Retrieval. In: *Proc. ACM Multimedia* (2001)
- [11] Tieu, K., Viola, P.: Boosting image retrieval. In: *Proc. IEEE Conf. Computer Vision Pattern Recognition*, pp. 228–235 (2003)
- [12] Zhou, X.S., Huang, T.S.: Relevance Feedback in image retrieval: A Comprehensive review. *Multimedia Systems* 8(6), 536–544 (2003)
- [13] Zhou, Z.-H., Chen, K.-J., Dai, H.-B.: Enhanced Relevance Feedback in Image Retrieval Using Unlabeled Data. *ACM Trans. on Informations Systems* 24(2), 219–244 (2006)
- [14] Ferecatu, M., Boujemaa, N., Crucianu, M.: Semantic interactive image retrieval combining visual and conceptual content description. *ACM Multimedia Systems Journal* 13(5-6), 309–322 (2008)
- [15] Hoi, S.C.H., Lyu, M.R., Jin, R.: A Unified Log-Based Relevance Feedback Scheme for Image Retrieval. *IEEE Trans. on Knowledge and Data Engineering* 18(4) (2006)
- [16] Tong, S., Chang, E.: Support vector machine active learning for image retrieval. In: *Proceedings of the 9th ACM International Conference on Multimedia*, pp. 107–118. ACM Press, New York (2001)
- [17] Tong, S., Koller, D.: Support vector machine active learning with applications to text classification. In: *Proceedings of ICML-2000, 17th International Conference on Machine Learning*, pp. 999–1006. Morgan Kaufmann, San Francisco (2000)
- [18] Kokare, M., Biswas, P.K., Chatterji, B.N.: Texture Image retrieval using New Rotated Complex Wavelet Filters. *IEEE Trans. on Systems, Man, and Cybernetics-Part B: Cybernetics* 35(6) (2005)
- [19] Hsu, C.-T., Li, C.-Y.: Relevance Feedback Using Generalized Bayesian Framework with Region Based optimization Learning. *IEEE Trans. on Image Processing* 14(10) (2005)

- [20] Ion, A.L., Stanescu, L., Burdescu, D.: Semantic Based Image Retrieval using Relevance Feedback. In: International Conference on Computer as a Tool, Warsaw, pp. 303–310 (2009)
- [21] Kingsbury, N.G.: Image processing with complex wavelet. *Phil. Trans. Roy. Soc. London A* 357, 2543–2560 (1999)
- [22] Kingsbury, N.G.: Complex wavelets for shift invariant analysis and filtering of signals. *J. App. Comput. Harmon. Anal.* 10(3), 234–253 (2001)
- [23] Selesnick, I., Baraniuk, R., Kingsbury, N.: The dual-tree complex wavelet transform. *IEEE Signal Process. Mag.* 22(6), 123–151 (2005)
- [24] Freund, Y., Schapire, R.E.: A decision –theoretic generalization of online learning and an application to boosting. *J. Comput. Syst. Sci.* 55(1), 119–139 (1997)

A New Method for Lossless Image Compression Using Recursive Crack Coding

T. Meyyappan¹, S.M. Thamarai², and N.M. Jeya Nachiaban³

^{1,2} Department of Computer Science and Engineering,
Alagappa University, Karaikudi – 630 003, India
meyslotus@yahoo.com, lotusmeys@yahoo.com

³ Department of Computer Science and Engineering,
Thiagarajar College of Engineering Madurai-9., India
nmjeyan2009@tce.edu

Abstract. Popular entropy coding methods for lossless compression of images depend on probability models. They start by predicting the model of the data. The accuracy of this prediction determines the optimality of the compression. These methods are very slow because of visiting the data (pixels) in left to right order. Parallel implementation of these methods is adopted by researchers to speed up the process. In this paper, the authors propose a new approach to image compression using crack coding. The novelty and better compression ratio of the method is due to its recursiveness in finding the variable-length entropy. The proposed method starts with the original image and develop crack codes in a recursive manner, marking the pixels visited earlier and expanding the entropy in four directions. The proposed method is experimented with sample bitmap images and results are encouraging. The method is implemented in uni-processor machine using C language source code.

Keywords: Contour, Crack Coding, Entropy, Lossless Compression, Image.

1 Introduction

Now a days network plays an important role in our life. It is hard to pass a day without sharing information with others. Transmission of images is also the need of the day. Transmission of images in their original form increases the time spent in network and we need to increase the bandwidth for fast transmission. On the other hand, compressed images which can be restored at the receiving end can very much reduce network overheads.

Compression of images[9] is concerned with storing them in a form that does not take up so much space as original. An enormous amount of data is produced when a 2-D light intensity function is sampled and quantized to create a digital image. Image compression addresses the problem of reducing the amount of data required to represent a digital image.

The underlying basis of the reduction process is the removal of redundant data. From a mathematical view point, this amounts to transforming a 2-D pixel array into a statistically uncorrelated data set[2]. The transformation is applied prior to storage or

transmission of the image. At some later time, the compressed image is decompressed to reconstruct the original image or an approximation to it.

2 Existing Methods

The four different approaches[3],[5] to compression are Statistical Compression, Spatial compression, Quantizing compression, Fractal compression. In spatial approach, image coding is based on the spatial relationship between pixels of predictably similar types. The method proposed in this paper employs spatial approach for compression.

Run-length encoding (RLE) is a very simple form of data compression in which runs of data (that is, sequences in which the same data value occurs in many consecutive data elements) are stored as a single data value and count, rather than as the original run. This is most useful on data that contains many such runs: for example, simple graphic images[8] such as icons, line drawings, and animations. It is not useful with files that don't have many runs as it could greatly increase the file size.

Huffman coding removes coding redundancy. Huffman's procedure creates the optimal code for a set of symbols and probabilities subject to the constraint that the symbols be coded one at a time. After the code has been created, coding and/or decoding is accomplished in the simple look-up table. When large number of symbols is to be coded, the construction of the optimal binary Huffman code is a difficult task.

Arithmetic coding (AC)[4] is a special kind of entropy coding. Arithmetic coding is a form of variable-length entropy encoding used in lossless data compression. Arithmetic coding differs from other forms of entropy encoding such as Huffman coding in that rather than separating the input into component symbols and replacing each with a code, arithmetic coding encodes the entire message into a single number.

Compression algorithms that use arithmetic coding start by determining a model of the data – basically a prediction of what patterns will be found in the symbols of the message. The more accurate this prediction is, the closer to optimality the output will be.

3 Image Model

A digitized image is described by an $N \times M$ matrix of pixel values are nonnegative scalars, that indicate the light intensity of the picture element at (i,j) represented by the pixel.

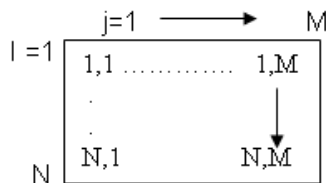
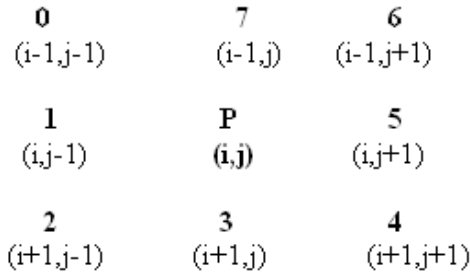


Fig. 1. Image Model

3.1 Connectivity

In many circumstances it is important to know whether two pixels are connected to each other, and there are two major rules[7] for deciding this. Consider a pixel called P, at row i and column j of an image; looking at a small region centered about this pixel, we can label the neighboring pixels with integers. Connectivity is illustrated below:



Two pixels are 4-adjacent if they are horizontal or vertical neighbors. The 4-adjacent pixels are said to be connected if they have the same pixel value.

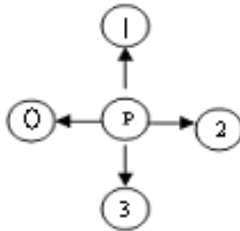


Fig. 2. 4-Connected Pixels

3.2 Contour Tracing with Crack Coding

Contour coding[4] has the effect of reducing the areas of pixels of the same grey levels to a set of contours that bound those areas. If the areas of same grey level are large with a simple edge, then the compression rate can be very good. In practice, it is best to make all contours circular[4], so that they return to the originating pixel - if necessary along the path that they have already traversed - and to identify the grey level that they lie on and enclose. 8-connected contour is known as chain coding and 4-connected contour is known as crack coding. In this paper, authors used crack coding and grey level of each contour is saved along with the contour.

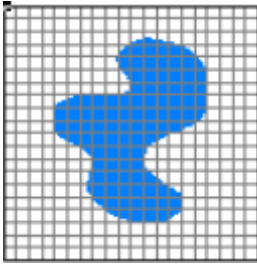


Fig. 3. (a) Original Image

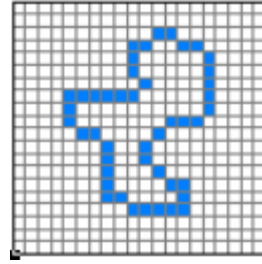


Fig. 3. (b) 4-Connected Contour

4 Proposed Method

The proposed method works with the original image as it is. It does not process the image in any way and transform the pixels of the image as in edge detection. It finds all the possible 4-connected contours and stores the 4-directions of the contour along with grey value being examined. The process is repeated with the help of a recursive procedure and marking all the pixels visited along the contour path. The marked pixels are eliminated for further examination of connected pixels. The four direction crack code values (0,1,2,3 consuming 3 bits per number) are packed into a byte and stored along with the grey value in output file. No loss of pixels[1] are observed in the proposed compression method. The following is the format of stored compressed image:

Row, Column, Grey-Value, 4-direction crack codes

4.1 Algorithm for Compressing Original Image

The following algorithm shows the sequence of steps to be followed to compress the original image.

- Step 1** Read an uncompressed image file[6]
- Step 2** Read number of rows n and columns m of the image from header
- Step 3** Separate pixels $P[n,m]$
- Step 4** For $i=1$ to n do 5
- Step 5** For $j=1$ to m do
 - Store $P[i,j]$ and its grey value g as beginning of the contour
 - Mark the pixel $P[i,j]$
 - Crack_Code(P,i,j,g)**
- Step 6** Write the header information and contour codes in another file.

Procedure Crack_Code(P,i,j,g)**Begin**

```

if(P[i, j-1] equal g) then store 0; Crack_Code(P,i, j-1,g);
else
if(P[i-1,j] equals g) then store 1; Crack_Code(P,i-1,j,g);
else
if(P[i,j+1] equals g) then store 2; Crack_Code(P,i, j+1,g);
else
if(P[i+1,j] equals g) then store 3; Crack_Code(P,i+1, j,g);
else
return;

```

End;**4.2 Algorithm for Restoration of Original Image from Compressed Image**

The following algorithm shows the sequence of steps to restore the original image from compressed image.

- Step 1** Open the compressed image file.
- Step 2** Read number of rows m and columns n of the image from header.
- Step 3** Initialize $P[n,m]$
- Step 4** Repeat steps 5 to 8 until all the crack coded contours are processed
- Step 5** Read starting coordinate position(i, j) and grey value g of next contour.
- Step 6** $P[i, j]=g$;
- Step 7** Read next crack code c ;
- Step 8** Replace_Pixel(P, i, j, g, c);
- Step 9** Write the header information and pixels $P[n,m]$ in another file.

Procedure Replace_Pixel(P,i, j,g,c)**Begin**

```

if(c equals 0) then store P[i, j-1]=g;
else
if(c equals 1) then store P[i-1, j]=g;
else
if(c equals 2) then store P[i, j+1]=g;
else
if(c equals 3) then store P[i+1, j]=g;
else
return;

```

End;

5 Results and Discussion

The authors have developed a package using C language code for the proposed compression and decompression methods. A set of sample bitmap images (both monochrome and color) are tested with the proposed method. The compression percentage varies from 50% to 90% for the samples. The percentage of compression is better for images with more number of similar grey values. No loss of pixels are found while restoring the original image. Original size and compressed size of the images and computation time are plotted.



Fig. 4. (a) Original Bitmap Image



Fig. 4. (b) IMAGE after Decompression

A sample content of the file which stores Starting Position of a pixel, Grey value and Crack Codes of its contour is shown below. The last value -1 marks the end of the contour.

```

0 0 200 3 2 2 2 2 2 3 0 0 -1
0 2 225 1 3 3 2 2 2 2 2 2 2 2 3 0 0 0 0 0 0 0 3 2 2 2 2 2 2 2 3 0 -1
0 7 175 3 2 2 2 2 2 3 2 2 2 2 -1
0 9 180 2 2 2 2 -1
1 7 190 2 2 2 2 -1
  
```

Table 1. Experimental Results

SAMPLE BITMAP IMAGES	ORIGINAL SIZE IN BYTES	COMPRESSED IMAGE SIZE IN BYTES	COMPRESSION PERCENTAGE	COMPUTATION TIME IN SECONDS	NUMBER OF BYTES DELETED
1	9108	6416	70.4	0.8	0
2	8036	6270	78.0	0.8	0
3	8415	6392	76.0	1.3	0
4	7698	6271	81.5	1	0

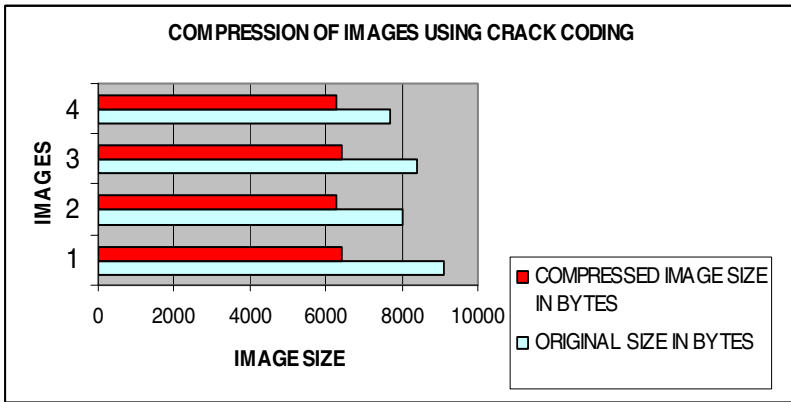


Fig. 5. Performance of the proposed compression method

6 Conclusion

The proposed method proves to be a lossless compression method. Program execution time, compression percentage and rate of information loss is measured for various images. Computation time for compression of an image is not directly proportional to the size of the image. It depends on the number of contours found in the image. Parallel processing methods may be adopted to speed up contour tracing operation. The next phase of the research work with 8-connected pixels (chain coding) is under progress.

References

1. Wu, X., Memon, N.: Context-based, adaptive, lossless image coding. *IEEE Trans. Commun.* 45, 437–444 (1997)
2. Ansari, R., Memon, N., Ceran, E.: Near-lossless image compression techniques. *J. Electron. Imaging* 7(3), 486–494 (1998)

3. Ekstrom, M.P.: Digital Image Processing Techniques (Computational Techniques). Academic Press, London (1984)
4. Low, A.: Introductory Computer Vision and Image Processing. McGraw-Hill Publishing Co., New York (1991)
5. Held, G., Marshall, T.R.: Data and Image Compression: Tools and Techniques. Wiley, Chichester (1996)
6. Miano, J.: Compressed Image File Formats: JPEG, PNG, GIF, XBM, BMP. ACM Press, New York (1999)
7. Sayood: Introduction to Data Compression, 2/e. Academic Press, London (2000)
8. Jahne, B.: Practical Handbook on Image Processing for Scientific and Technical Applications. CRC Press, Boca Raton (2004)
9. Parker, J.R.: Algorithms for Image Processing and Computer Vision. Wiley, Chichester (2010)

A Study on Block Matching Algorithms and Gradient Based Method for Motion Estimation in Video Compression

Deepa Mary Thomas

M.Tech Scholar, Department of Computer Science and Engineering
Rajagiri School of Engineering and Technology,
Rajagiri Valley Kakkanadu, Kerala, India
marythomasdeepa@gmail.com

Abstract. This paper mainly focuses on the classification of motion estimation algorithms used for video compression. Motion Estimation (ME) algorithms vary with respect to the a priori information and constraints they employ, as well as the method of computation they use to obtain the estimate. The classifications for ME algorithms are based on feature/region matching, gradient based methods, spatiotemporal energy methods, deterministic model based methods and stochastic model based methods. This paper focuses more on block matching algorithms (BMA), (used for motion estimation in video compression) which comes under feature/region matching and gradient based methods. In effect this paper compares 7 different types of block matching algorithms. In addition to the above comparisons a survey on the Shape-Adaptive Motion Estimation Algorithm for MPEG-4 Video Coding is also incorporated. The Shape-Adaptive Motion Estimation algorithm is based on the gradient based motion estimation method. It is one of the latest algorithm coming under gradient based method. The algorithms that are compared and discussed in this paper are widely accepted by the video compressing community. More over these algorithms are used for implementing various standards, which ranges from MPEG1 / H.261 to MPEG4 / H.263 and H.264/AVC. A very brief introduction to the entire flow of video compression is also presented in this paper.

Keywords: Block matching, Gradient based method, Motion estimation, MPEG, H.261, H.263, H.264/AVC, Video compression.

1 Introduction

With the onset of the multimedia age and the wide use of Internet, video storage on CD or DVD and streaming video has become highly popular. The digital video application has gained wide appeal in mobile terminals. Examples relating to this are personal digital assistance and cellular phones. But a major problem still remains in a video is the high requirement for bandwidth. For example a typical system have to send large number of individual frames per second to create

an effect or illusion of a moving picture. With the limited storage and bandwidth capacity, this data (raw video data) must be compressed to a transportable size. For this reason, several standards for compression of the video have been developed. Many video compression standards were already in vogue. Digital video coding has gradually become very popular since the 1990s when MPEG-1 first emerged. The video coding achieves higher data compression rates and also it eliminates the need of high bandwidth requirements. The important aspect in the video coding is that without any significant loss of subjective picture quality, the video coding achieves higher data compression rates. The MPEG standards has the advantage of having lesser bandwidth for data transmission. Currently MPEG-4 has become the dominant video codec algorithm for streaming and distribution of video contents across wireless media at low bandwidth. The latest approved video compression standard, known as MPEG-4 Part 10 AVC or H.264 AVC (AVC - Advanced Video Coding), has shown a 50 percentage compression improvement when compared to the previous standards. This compression improvement comes at the cost of increase in computational complexity, which results in higher computation power dissipation. One such important cause for this computation power dissipation is Motion Estimation (ME). Motion estimation techniques form the core of video compression and video processing applications. Majority of the time is used for ME in video compression process. This paper will discuss more on the block matching ME algorithms coming in feature/region matching class. The entire flow for video compression is explained in section 2. Motion Estimation and a brief classification of ME algorithms will be present in section 3 and 4. Then the coming section, section 5 will analyze different block matching algorithms that range from the very basic Exhaustive Search to the latest GA based 4GS ME algorithm. Shape-Adaptive Motion Estimation Algorithm for MPEG-4 Video Coding is included in section 6. In section 7 comparison among different block matching algorithms were discussed. Summary and references will be provided in the last sections.

2 Video Compression Process

Video compression means to reduce data redundancy and irrelevancy. The sources of data redundancy are of spatial, temporal and color space (statistical correlation between frames). In spatial redundancy the nearby pixels are often correlated (as in still images). The temporal redundancy deals with the adjacent frames. The irrelevancy deals with perceptually unimportant information. The main addition over image compression, in video compression is to exploit the temporal redundancy. In video compression process we can predict the current frame based on previously coded frames (that is reference frame). Three types of coded frames are included in video compression (I-frame, P-frame and B-frame). The two blocks that is the encoder and the decoder will work as follows. The encoding side calculates the motion in the current frame and compares it with a previous frame. A motion compensated image for the current frame is then created (motion compensated image is built of blocks of image from the

previous frame). The motion vectors for blocks used for motion estimation are transmitted. In addition to that the difference of the compensated image with the current frame is also JPEG encoded and sent. After that the encoded image that is sent is then decoded by the encoder. This image is then used as a reference frame for the subsequent or coming frames. In the decoder side the decoder will reverse the entire process and then creates a full frame. Generally speaking, video compression can be defined as a technology for transforming video signals. The main focus in video compression is to retain the original quality under a number of constraints. For example, storage constraint, time delay constraint or computation power constraint.

3 Motion Estimation (ME)

Before dealing with ME or to understand what is Motion Estimation, it is essential that we have to know about what MPEG is, of which motion estimation is a part. Moving Pictures Expert Group (MPEG) is actually a body used for developing a suitable encoding scheme or what are called standards. It is used for transmitting moving pictures and sound over various broadcast links. The various standards of MPEG released so far are: MPEG-1, MPEG-2 and MPEG-4. Under MPEG-4 itself there came many parts, of which the latest international video coding standard is H.264 Advanced Video Coding AVC or MPEG-4 part 10 AVC. The most computationally expensive and difficult operation in the entire compression process is Motion Estimation. Motion Estimation is one of the most critical modules in a typical digital video encoder. ME can be defined as part of 'inter coding technique'. The Motion estimation process extracts the entire motion information from the video sequence. Inter coding refers to a mechanism of finding co-relation between two frames (still images), which are not far away from each other as far as the order of occurrence is concerned, one frame is called the reference frame and the other frame called the current frame, and then encoding the information which is a function of this co-relation instead of the frame itself. In ME each block in a frame is represented by a motion vector that represents the displacement of the block with respect to another highly correlated block in a previously encoded frame [1],[2],[3]. The concept of Motion Compensation (MC) technique is that to provide a better prediction of the current frame, the encoder uses the motion model and informations to move the contents of the reference frame. This process is known as motion compensation (MC) [4],[5],[6],[7], and the prediction produced for this purpose is called the motion compensated prediction (MCP) or the displaced-frame (DF). The difference of ME from motion compensation is that, the ME detect the movement of objects in an image sequence and it will try to obtain the motion vectors representing the estimated motion. Apart from ME the motion compensation techniques uses the knowledge of object motion so obtained in order to achieve data compression. Both Motion Estimation (ME) [8] and Motion Compensation (MC) techniques are used to reduce the temporal redundancy between successive frames, which is in the time domain. When ME is performed by an MPEG-2 encoder it groups the pixels

into $16 * 16$ macro blocks. The MPEG-4 AVC encoders can further divide these macro blocks into small partitions (as small as $4 * 4$).

4 Classification of ME Algorithms

The ME algorithms are classified into feature/region matching, gradient based methods, spatiotemporal energy methods, deterministic model based methods and stochastic model based methods. The ME algorithms that are most commonly used in video image processing belong to either the feature/region matching or the gradient based methods classes. In feature/region matching the motion is estimated by correlating/matching features (e.g., edges) or regional intensities (e.g., block of pixels) from one frame to another. The block matching algorithms and phase correlation methods will come under this class. The Gradient-Based Methods estimate the motion by using spatial and temporal changes (gradients) of the image intensity distribution as well as the displacement vector field. The methods coming under this class are Pel-Recursive methods (which derive motion vector for each pixel) , Netravali-Robbins, Walker-Rao, Wiener estimation based and the Horn Schunck algorithm. Both the matching algorithms and the gradient based methods comes under time-domain algorithms. This paper evaluates the fundamental block matching algorithms from the mid-1980s up to the recent fast block matching algorithms of year 2010.

5 Block Matching Algorithms (BMA)

The most common ME method is the block matching algorithm (BMA). Block Matching Algorithm (BMA) is the most popular and famous ME algorithm. Instead of individual pixels, BMA calculates the motion vector for an entire block of pixels. The same motion vector is applicable to all the pixels in the block. This reduces computational requirements and also results in a more accurate motion vector since the objects are typically a cluster of pixels. In BMA, the current frame is divided into a number of pixel blocks (matrix of macro blocks) each of which consists of luminance and chrominance blocks. For coding efficiency, ME is performed only on the luminance block. The motion estimation is performed independently for each pixel block (that is for luminance block). For doing ME identify a pixel block from the reference frame that best matches with the current block, and whose motion is being estimated. The matching of one macro block with another block is based on the output of a cost function (error criteria). Which macro block matches closest to the current block is computed by finding the macro block that results in the least cost. The most commonly used error criteria are the mean square error (MSE) or the sum of square error (SSE) and the minimum absolute difference (MAD) or the sum of absolute difference (SAD). The SSE gives a more accurate block matching, however it requires more computations. The SAD provides fairly good match at lower computational requirement and because of this SAD is widely used for block matching. Some of the issues coming in block matching is matching criterion, search procedure and

block size. Peak-Signal-to-Noise-Ratio (PSNR) characterizes the motion compensated image. This image is created by using MVs and macro blocks from the reference frame.

5.1 Exhaustive Search (ES)

This algorithm is also known as Full Search (FS) Motion Estimation algorithm. It is the most computationally expensive block matching algorithm of all the existing algorithms. Full Search Block Matching algorithm evaluates every possible pixel blocks in the search region (search area). By evaluating each and every possible pixel blocks in the search area this algorithm can generate the best block matching motion vector (MV). As a result of which it finds the best possible match. Also it returns the highest PSNR amongst any block matching algorithm. The FS calculates the SAD value at each possible pixel location in the search region. For analyzing each block in the current frame, each and every candidate block that resides within the search window (or search region) in the referenced frame is searched. The motion vector (MV) is then calculated and it points to the block with the lowest distortion (minimum distortion value) in the reference frame. This type of BMA can give least possible residue for video compression. But, the required computations are exceedingly very high due to the large amount of candidates present in each blocks to get evaluated. It does not guarantee consistent motion vectors required for video processing applications. Other fast block matching algorithms try to achieve the same PSNR value produced by FS, by doing little computation as possible. The Full Search results in global minima [8] whereas other fast BMA algorithms results in selecting a candidate corresponding to local minima. It is impractical or less efficient in terms of computational complexity or runtime. The three-step search (TSS) [2], new three-step search (NTSS) [10], four-step search (4SS) [11], diamond search (DS) [9] algorithms are amongst the class of fast search methods used. They reduce the NSPs (number of search points) in the process of block motion estimation.

5.2 N-Step Search

The N-Step Search (more specifically Three Step Search (TSS)) is one of the most commonly used fast search algorithms. The steps of the algorithm are summarized as follows: Initializing: Let S be the step size. 1) An initial step size (S) is chosen. This S is usually half of the maximum motion displacement. Then for comparison the eight blocks at a distance of the step size (S) from the centre point are chosen. 2) After that the step size (S) is halved and the centre point is moved to the block with the minimum distortion value. 3) Repeat steps 1 and 2 until the step size becomes one. The TSS, NTSS, 4SS all comes under the classification, fixed search area (which is under search area sub sampling).

5.3 Three Step Search (TSS)

This is one of the earliest attempts at fast block matching algorithms. The hunt for this type of algorithm were started in mid 1980s. The Three Step Search

(TSS) was introduced by Koga et al in 1981[19]. The main attraction of TSS over the preceding algorithms is that they are simple and robust in nature. Because of the near optimal performance the TSS became very popular for low bit-rate video applications. The searching pattern for TSS is that it searches for the best MVs in a course to fine search patterns. The idea behind TSS is that in every macro block the error surface due to motion is unimodal. One of the characteristics for this surface is that the weights generated by the cost function will increase monotonically from the global minimum. The TSS algorithm is described in 3 steps. In the first step it starts with the search location at the center. Furthermore it sets the step size $S = 4$, for a usual search parameter p value of 7. After setting up the S value and the corresponding p value it then searches around the location $(0,0)$ for the 8 locations, $+ S$ pixels or $- S$ pixels. From these 9 locations searched so far it picks the pixel giving the least cost value and makes it the new search origin. Secondly after getting the new search origin it sets the new step size value as $S = S/2$, and for until $S = 1$ this process repeats similar search for two more iterations. At that point (when $S=1$) it finds the location with the least cost function. The macro block at that location or the location with least cost function is the best match. In the final step the calculated motion vector (MV) is then saved for transmission. TSS gives an improved flat reduction in the computation process by a factor of 9. The main advantage for TSS is that the number of candidates evaluated during the three-step search (TSS) procedure is very less when compared to the full search (FS) algorithm. TSS is not efficient to catch small motions appearing in stationary as well as quasi-stationary blocks.

5.4 New Three Step Search (NTSS)

After the development of TSS, in 1994 the New Three Step Search came into existence. Based on the characteristic of center-biased motion vector distribution, a new three step search (NTSS) algorithm for improving the performance of the TSS on the estimation of small motions was developed. Like TSS the NTSS is also used for fast ME. NTSS [10] provides much improvement over TSS results. The NTSS was considered as one of the first widely accepted fast algorithms. This algorithm is frequently used for implementing MPEG 1 and H.261 (that is for developing the earlier standards). The TSS uses a uniformly allocated checking point pattern for motion detection and one of the pitfall for TSS is that it is very prone to missing small motions. In NTSS, first of all the cost functions to be determined. In the first step using one of the cost function, 16 points are checked in addition to the search origin (point) for minimum weight. Totally there will be 17 points if we consider the origin also. The additional search locations, that is 8 are a distance of $S = 4$ away (similar to TSS) and the other 8 points are at $S = 1$ away from the search origin. If the lowest (minimum cost function value) cost is at the origin then the search is stopped suddenly. Then the motion vector (MV) is set as $(0, 0)$. Furthermore it also has the worst-case scenario of 33 locations to check. The search pattern used in NTSS for each step is fixed. This algorithm does not use any kind of threshold operations. This

algorithm is adaptive. NTSS resembles like the original TSS and it is simple in nature. Comparing with TSS it works better in terms of motion compensation module. NTSS is very much compatible with the TSS in terms of computational complexity. The new NTSS algorithm differs from TSS in 2 aspects. They are 1) used a center-biased checking point pattern in its first step, in order to make the search adaptive with the distribution of MV. 2) added a halfway-stop technique for stationary blocks or quasi-stationary blocks.

5.5 Four Step Search (4SS)

Another algorithm that also adopts the center-biased pattern feature is 4SS. The 4SS performs better than the well-known three-step search (TSS) and has similar performance to the new three-step search (NTSS) in terms of motion compensation errors. The 4SS also reduces the worst-case computational requirement from 33 to 27 search points and the average computational requirement from 21 to 19 search points as compared with NTSS. Similar to NTSS, 4SS has a halfway stop provision. 4SS uses a fixed type of pattern size, say $S=2$ for the first step. 4SS looks at 9 locations in a 5×5 window. 3SS gives importance to 9×9 window. The 4SS algorithm reduces the number of search points (NSP) very largely when compared to TSS and we can say that it is more robust.

5.6 Diamond Search (DS)

Diamond Search [9] algorithm is exactly the same as 4SS algorithm. Only change that is to be made is that the search point pattern is changed from a square pattern to a diamond pattern. There is no limit put on the number of steps that the algorithm can take. In the diamond search (DS) algorithm the step size (S) is always set to one. The algorithm is designed to reduce the computational complexity or runtime, even more than the other fast search algorithms. As an output an acceptable reconstructed PSNR is returned. For each block in the current frame, the blocks in the reference frame, making a diamond pattern, are searched for the block with the minimal distortion. The DS algorithm uses two different types of fixed patterns. They are Large Diamond Search Pattern (LDSP) and Small Diamond Search Pattern (SDSP). In DS algorithm the first step uses LDSP. If the least weight (cost value) is at the center location then a jump is made and the search window size is reduced from 5×5 to 3×3 . The last step uses SDSP around the new search origin. The attraction of DS over other fast algorithms is that as the search pattern is neither too small nor too large and the fact that there is no limit drawn to the number of steps, this algorithm can find global minimum very accurately. The end result should give a PSNR value close to that of ES or FS. At the same time the computational expense should be significantly less. The DS algorithms main highlight is that this will consistently performs well for the video images with large range of motion content. It also outperforms one of the well-known TSS algorithm. It is having a close performance with NTSS. It reduces the complexity of computation approximately by 20 - 25 percent.

5.7 Four Step Genetic Search (4GS)

This algorithm, 4GS is used for block matching purposes. The main contribution to this algorithm is given from the genetic algorithm (GA) side. This 4GS ME algorithm combines both the features of GA and 4SS. The method has shown a similar performance rate when compared to FS in terms of one of the cost function, that is MSE. The criteria for the number of search points (NSP) is a problem. This algorithm takes NSP equal to that of 14 percent of FS and the rate of NSP when compared with 3SS is very similar. Another ratio called the speed up ratio between this algorithm and FS is 5.6 times. 4GS need more search points than 3SS and 4SS. Computation time of 4GS is much less than FS. They can save upto 80 to 85 percent of computation time. The performance of GA based BMA depends on the nature of image sequences. This algorithm shows some regularity and so it is suitable for hardware implementations. It is suitable for standards like H.261, MPEG-1 and MPEG-2.

6 Gradient Based Method

In the block based techniques the MV can be determined based on the variations in the pixel intensities. The best matching MV is considered as the pixel with less intensity difference between the current frame and the reference frame. But in the gradient based techniques, the spatiotemporal derivatives of pixel intensities is calculated to determine the MV values. The total derivative of the I function should be zero for each position in the image. It should be made as zero every time and the I function is the image intensity function. During the search process, there comes a problem. That is we have to find the MV for the current block for time instance. Therefore the SAD value between the current block and the matching block at that time instance is minimized.

6.1 Shape Adaptive Motion Estimation Algorithm

The algorithms are meant for MPEG-4 video coding purposes. So one of the latest gradient based ME algorithms that works well on the shape motion prediction is this Shape-Adaptive Motion Estimation Algorithm. This algorithm takes the advantage of the correlation function. This function is between the neighboring Binary Alpha Blocks (BABs). This correlation between blocks is used to match with the Mpeg-4 shape coding scenarios and to speed up the estimation process. The PSNR values and the computation time achieved by this algorithm seems to be better than those obtained by other ME techniques. The shape information is used to reduce the number of search point (NSP) per macroblock. This algorithm combines the motion estimation for shape coding (the shape motion estimation is applied for boundaries only) with the motion estimation techniques for texture (texture motion estimation is for the opaque macro-blocks). For this algorithm [12] the advantage found out is that it takes minimum number of iterations for the algorithm and also the computation time

needed is very less. Again the algorithm exploits the optical flow principle. It uses a recursive ME which is a less complex method. This recursive ME is used to calculate the dense displacement fields. The limitation highlighted is that it shows degradation in the picture quality. The algorithm is described as follows. The algorithm is divided into 2 steps. The first step used is the Block Recursive Search (BRS). In this step 4 candidate MV, that is 3 spatial and 1 temporal are checked for the exact or actual block with the help of a recursive block matching. The second step involved is the Pixel Recursive Search (PRS). For this step the selected vector is adjusted by a gradient method to find the suitable approximation. In BRS, the video frames are scanned (raster scan is used) from top left to the bottom right to evaluate the candidate vectors depending on the type of the macro blocks. The macro blocks can be opaque, transparent or boundary. Efficiency is great for this algorithm. This algorithm also gives a PSNR value similar to that of FS.

7 Comparison of the Above Discussed BMA Algorithms

The earliest technique Full Search, which comes as a brute force technique shows the searching pattern as, here each and every block is matched with the current block to find the best suitable match. The technique provides accuracy and it is simple in nature. But this search is not so efficient because it suffers from high computational cost. FS provides good quality images. So other fast and highly efficient block matching algorithms came into existence. For each ME algorithms to survive there is a well known trade off between the efficiency and the quality (of images). FS provides good quality images but it lacks low computational complexity. So by considering this reason many new fast algorithms came. The N-Step Search came and there the most popular algorithm is TSS. One problem that occurs with the TSS is that for the first step it uses a uniformly allocated checking point pattern. By using this uniformly allocated checking point patterns it is difficult to identify the small motion estimations. The advantage of TSS is that it reduces the amount of candidate vectors needed for evaluation. The NTSS reduces the computation cost by using half way stop provisions. The 4SS, DS come pretty close to the PSNR results of FS. All the fast algorithms like TSS, NTSS, 4SS and DS they tried to achieve the same PSNR value as that of FS, but they consider only the most needed motion vectors. The shape adaptive ME which uses the gradient based technique yields a PSNR similar to that of FS. But image quality is not so good. Lastly the gradient based method described the shape adaptive ME is used for implementing many architectures.

8 Conclusion

In this paper we presented an overview of what a video compression process is and how we can implement the ME techniques. Furthermore we briefly introduced BMA in video coding. Based on the BMA techniques, we discussed about

8 different block matching algorithms that range from the very basic Exhaustive Search to the recent fast 4GS ME algorithms. Usually the ME is the quite computationally intensive and expensive step in the video compression process. It can consume up to 75-80 percent of the computational power of the encoder if the FS is used by exhaustively evaluating all the possible candidate blocks within the search window. As a consequence, the computation of video coding is greatly reduced with pattern based block motion estimation. We also stated the difference between the block based ME techniques and the gradient based methods. Finally a shape adaptive binary block matching ME is discussed and its performance is good and also the computational complexity increase can be accepted.

References

1. Richardson, I.E.G.: Video Codec Design, ch. 4, 5, 6. John Wiley & Sons Ltd., West Sussex (2002)
2. Turaga, D., Alkanhal, M.: Search Algorithms for Block Matching in Motion Estimation (1998)
3. Gibson, A.: The H.264 Video Coding Standard (August 2002)
4. Koga, T., Iinuma, K., Hirano, A., Iijima, Y., Ishiguro, T.: Motion-compensated interframe coding for video conferencing. In: Proc. NTC 1981, New Orleans, LA, pp. C9.6.1–C9.6.5 (December 1981)
5. Jain, J.R., Jain, A.K.: Displacement measurement and its application in interframe image coding. *IEEE Trans. Commun.* COM-29, 1799–1808 (1981)
6. Srinivasan, R., Rao, K.R.: Predictive coding based on efficient motion estimation. *IEEE Trans. Commun.* COM-33, 888–896 (1985)
7. Kappagantula, S., Rao, K.R.: Motion compensated interframe image prediction. *IEEE Trans. Commun.* COM-33, 1011–1015 (1985)
8. Motion Estimation Techniques in Video Processing, By Milind Phadtare System Architect, Video NXP Semiconductors India Pvt. Ltd. (August 2007)
9. Zhu, S., Ma, K.-K.: A New Diamond Search Algorithm for Fast Block-Matching Motion Estimation. *IEEE Trans. Image Processing* 9(2), 287–290 (2000)
10. Li, R., Zeng, B., Liou, M.L.: A New Three-Step Search Algorithm for Block Motion Estimation. *IEEE Trans. Circuits And Systems For Video Technology* 4(4), 438–442 (1994)
11. Po, L.-M., Ma, W.-C.: A Novel Four-Step Search Algorithm for Fast Block Motion Estimation. *IEEE Trans. Circuits And Systems For Video Technology* 6(3), 313–317 (1996)
12. Donghoon, Y., Sung Kyu, J., Jong Beom, R.: Fast motion estimation for shape coding in MPEG-4. *IEEE Trans. Circuits Syst. Video Techn.* 13(4), 358–363 (2003)

Automatic, Robust Face Detection and Recognition System for Surveillance and Security Using LabVIEW (sCUBE)

C. Lakshmi Deepika, M. Alagappan, A. Kandaswamy, H. Wassim Ferose, and R. Arun

Department of Biomedical engineering,
PSG College of Technology,
Coimbatore -641004, India

Abstract. The automatic, high end surveillance systems are of immense need in the wake of the emerging security problems faced in today's world. Most of the high end systems use current trends in technology but often prove to be costly which make them un-affordable for the common people. Thus there is an urge to develop a fully functional, high end, continuous surveillance system which has an error free monitoring and also cost effective. Thus we have taken up the challenge of developing a low cost, real time face detection and face recognition system which can provide automatic, robust, unmanned Surveillance and Security at critical points. The system was successfully installed and the efficiency of the overall system was tested.

1 Introduction

The application we have developed is a Security System, which uses Face Extraction and Recognition as its underlying principle. Face is one of the qualifying biometric identifier, which can be used to identify an individual effectively. Since the face biometric possesses the merits of high accuracy and low intrusiveness, it has drawn attention from various researchers especially in areas of security. An automatic system is very useful when due to the inherent nature of a human operator such as boredom or reduced alertness levels; he is not able to effectively scrutinize the voluminous videotape produced by the surveillance cameras. This proposed system can be used to make the security at the important places more effective in cases such as authenticating authorized persons, identifying in advance the visiting of VIPs, for detection of a criminal, for protection of undesirable property loss and to save valuable lives by defending or guarding from attack.

We hence propose to solve the challenge by providing visual facility to a computer and convert it to a multi functionality system which will continuously monitor a given critical entry point or area using a high resolution camera. The system will be able to detect human face from the camera output, recognize it from a database and allow the person if authorized or indicate an emergency if unauthorized. The image processing algorithms we have developed for face detection and recognition are computationally

complex and require a dynamic image processing tool to adapt it. The Vision Assistant, DAQ Assistant and Serial Port Communication in LabVIEW made it possible to successfully implement our idea in the proposed system.

The graphical system design of our system enhances the programming ability of the programmers. Its versatile design and the application enable the maximum dynamism in developing the complex application such as ours, in a better way rather than the text based programming languages. It reduces the development time of applications as well as increases the interest of developing new algorithms. The utmost user-friendly structures creates an environment to identify the application easily even by the idle man. The attractive visualization of the tools induces new innovations in the design during development and enriches the effective handling capability in the real time implementations.

There are several approaches to face detection and recognition in literature. Kanade presented an automatic feature extraction method based on the ratios of distances between facial features and reported a recognition rate of about 45-75% with a database of 20 people [1]. Brunelli and Poggio compute a set of geometrical features such as nose width and length, mouth position and chin shape. They were able to obtain 90% recognition with a database of 40 people [2]. Turk and Pentland proposed a face recognition scheme in which the face images are projected onto the principal components of the original set of training images [3]. This method also called Principal Component Analysis (PCA) later became a defacto standard in the field of pattern recognition and also as benchmark for testing new methods in research on pattern recognition. We propose to use PCA in our real-time system.

2 Image Acquisition

A camera is mounted at a vantage point at the entrance of the area under surveillance. The entry of a person into the area is detected by using a laser beam. When the laser

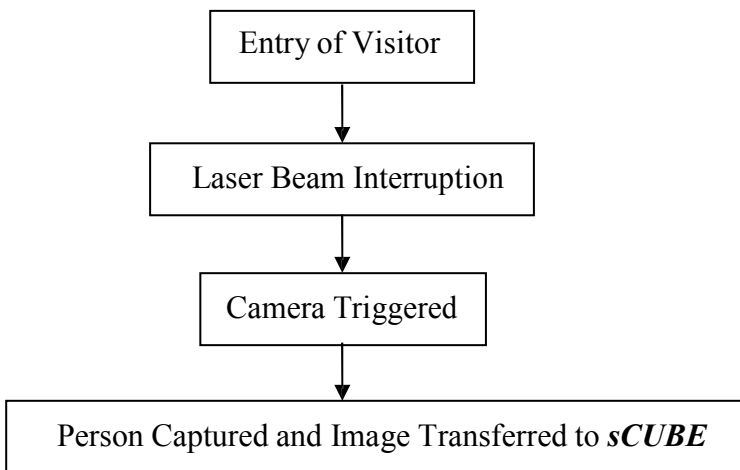


Fig. 1. Steps involved in Image Acquisition

beam is cut due to the person crossing by, it produces a software trigger to the camera, which creates a snapshot of the area under its focus. This photograph will have the person as well as the background and any other objects in it. It is transferred to the PC in which our LabVIEW application is installed. The Real time Image Acquisition facility in LabVIEW package helped us to simultaneously process the photograph as and when the person was captured in the area under surveillance.

3 Face Detection

This can also be called Face Segmentation where the facial images have to be segmented from the background. This task is seemingly effortless when it is done

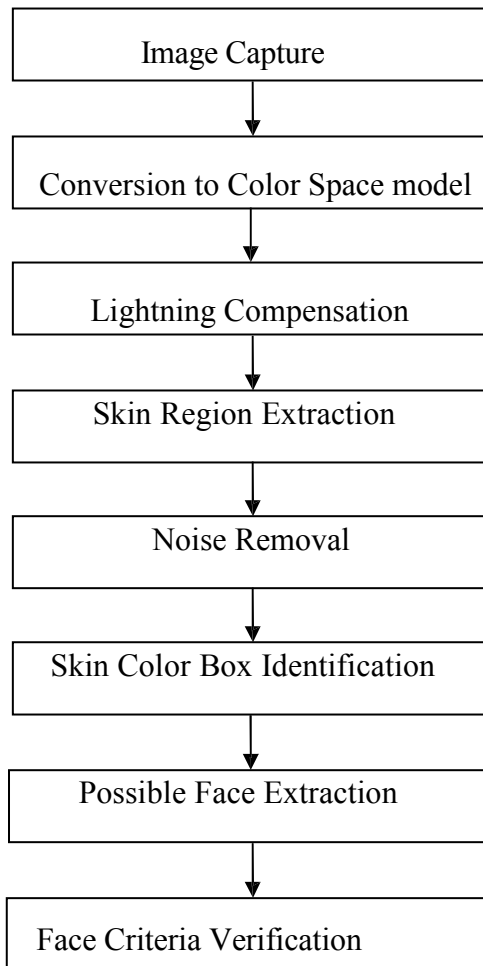


Fig. 2. Steps involved in Face Detection

by a human; however the underlying computations within the human visual systems are very complex. A full-fledged Surveillance and Security system has to detect locate and segregate faces, which will then be given as input to the Face Recognition system. According to the literature survey, it is found that the different human skin colors from different races, falls in a compact region of the color spaces. Hence, the image obtained from the camera is first transformed to the color space model namely the YCbCr model which gives the best results for skin pixel detection [4]. The characteristics of the face can be changed due to unexpected conditions such as shadows or change in lighting. Hence the lighting compensation is provided, the skin regions are extracted and the noise components are removed. Next the skin regions are labeled and identified and the labeled regions are tested for face criteria. Finally the possible faces are extracted.

Hence the automation of this task is a complex one, but the Vision Development Module of LabVIEW package with its vast library of powerful operators helped us to implement our complex algorithms magnificently.

4 Face Recognition

It is basically a machine learning or pattern classification problem. The extracted face image is converted to a set of features called the feature vector and compared with the facial database. This is also a quite challenging task since the identification of the face image has to be done in spite of variations due to variable illumination, facial expression and decoration (eg., glasses). It may also be required to identify the test face image from a large gallery of facial images, which is also a skilful development task.

To perform the face recognition, here a statistical tool namely Principal Component Analysis is used [5-7]. The power of PCA lies in its ability to convert a large dimensional data space (the image database) to a smaller dimensional feature space, so that the presence or absence of the test image in the database can be found with less number of computations. In this method, every face image, which is a 2D matrix, is converted into 1D arrays each of dimension N . The M arrays corresponding to M face images, that are arranged as a 2D array to form the data set and the empirical mean of all the N columns are calculated. The mean centered 2D array is formed by finding the deviation of every column from its mean. The Covariance matrix is then constructed. The Eigenvectors and Eigen values of the matrix are found and arranged according to their values. The largest eigen values are the principal components or eigen faces, which forms the basis for the large image space. The input test face is projected into this eigen face space and using Euclidean Distance metric, the nearest matching face image is retrieved from the database

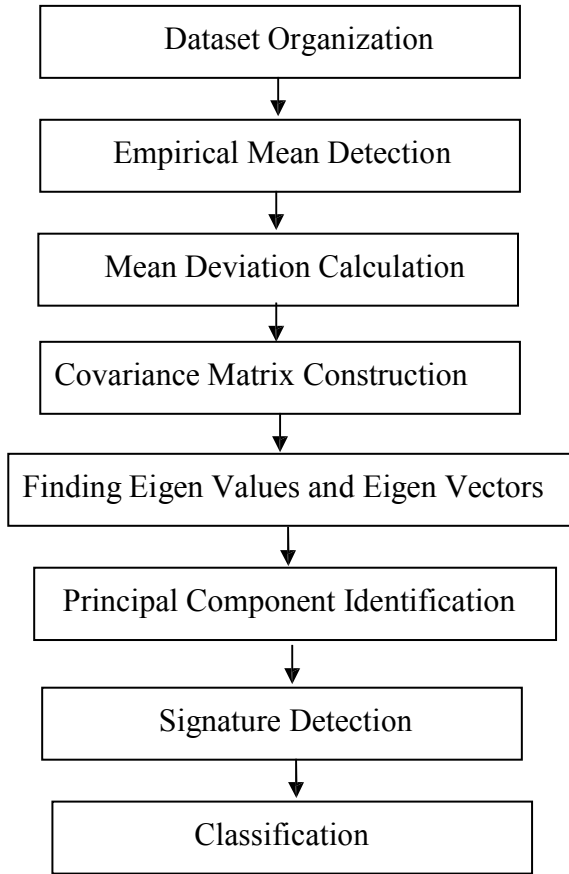


Fig. 3. Steps involved in Face Detection

5 Implementation Using LabVIEW

The implementation of the face recognition and detection system is explained as four modules:

A. Image Acquisition

The camera used here is an NI Compatible Scout camera and is triggered by the sensor that is mounted in front of the door, which is activated during the nearby visit of the person. Then the immediate image taken by the camera is processed by the face detection section.

B. Face Detection

- The captured RGB image is converted into YCbCr color space model using user defined functions.

- The lighting compensation is carried on to the test image using statistical formulas to isolate the background color.
- The skin regions are extracted using functions available in Vision and Motion tools.
- The noise is removed from the extracted skin region using the smoothing filter.
- The skin color box is identified using Blob analysis.
- The face criteria are implemented using the functions available in Vision and Motion tools and programming tools.
- The possible face region is detected after checking the face criteria.

C. Face Recognition

- Every face image which is in the database is converted into a 1D array of N elements where N is the number of pixels in the image. Every array corresponding to single images are then appended to form a 2D array to form the dataset.
- The empirical mean of every column is found.
- The deviation of every column from its mean is then found. The covariance matrix is then constructed.
- The eigen values and eigenvectors of this matrix are calculated and sorted.
- The largest eigen values are taken to form the basis of the eigen face space.
- The test face is now projected to this eigen face space.
- The nearest matching image is retrieved from the database using Euclidean distance metric which gives the image of the person passing by.

All the above steps are performed using the functions available in Programming tools, Mathematics tools, and Vision and Motion tools.

D. Database Management

Here are some of the features which we have implemented in sCUBE using functions available in programming tools.

- View Database – To view the complete list of authorized visitors and their personal details.
- Appending to Database – To add the new record into the database.
- Remove from Database - To remove the specific entry from the database.
- Specific Search – To search for a specific visitor and his visiting times.
- General Search – To search for all the visitors and their visiting times.
- View Sent Mail - To view the list of E-Mails sent during the visit of unauthorised persons.
- Maintenance of image database for unauthorized visitors.

6 Results and Discussion

A. Face Detection



Fig. 4. Face Detection Performed on a test image taken using our camera

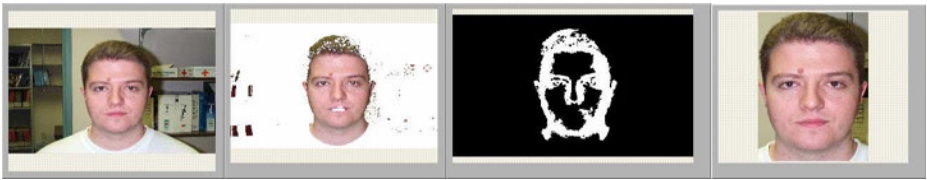


Fig. 5. Face Detection Performed on a test image available in Internet database

B. Face Recognition



Fig. 6. Face Recognition performed for the corresponding Detected Image



Fig. 7. Face Recognition performed for the corresponding Detected Image

To verify the performance of our face recognition module we have used freely downloadable databases from the internet (NLPR database with 450 images, AR database with 4000 images, ORL database with 400 images), we have also created our own database with 200 images similar to the above public database, under conditions of varying pose, facial expression and illumination. The results we obtained by using the ORL database is given in Table-1:

S.No	Number of Principal Components	Recognition Rate (using Euclidean Distance)
1	50	73.27
2	60	73.59
3	70	73.95
4	80	73.84

We were able to obtain an accuracy rate of 73.95%. However we were able to obtain about 90.18% accuracy rate when we applied our algorithm to our own database consisting of images with very less variation in pose, expression and illumination compared to the ORL database.

C. Database Management

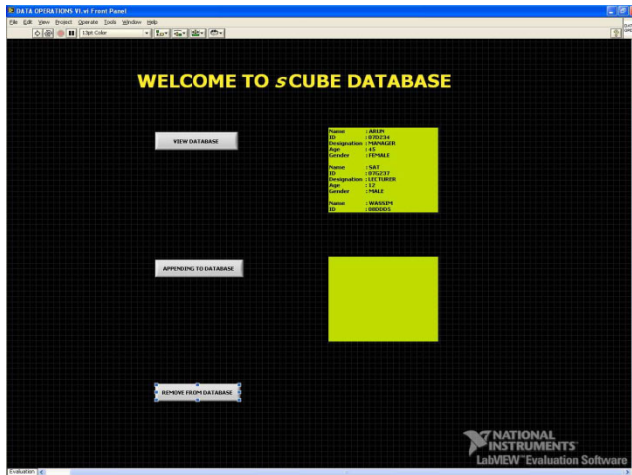


Fig. 8. Database Management – View Database, Appending to Database, Remove from Database

D. Other Features

- Voice Alert – To welcome the authorized entries.
- Alarm – Raised during the entry of unauthorized person . Opening/ Closing of doors – Automated for the authorized entry.
- SMS Alert – Sent to the control room during the unauthorized entry.

The following is the front panel of our proposed system

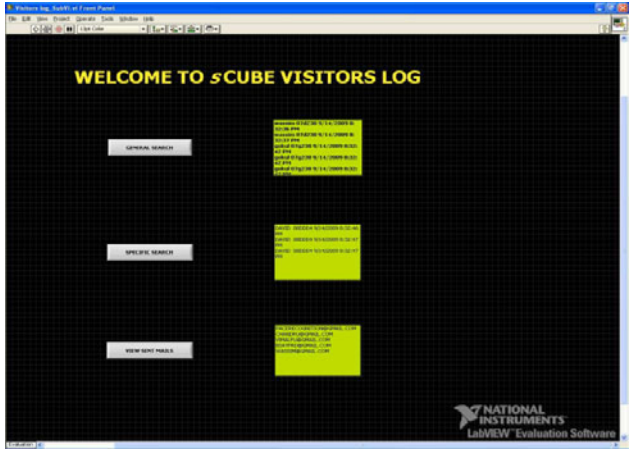


Fig. 9. Database Management – General Search, Specific Search, View Sent Mail

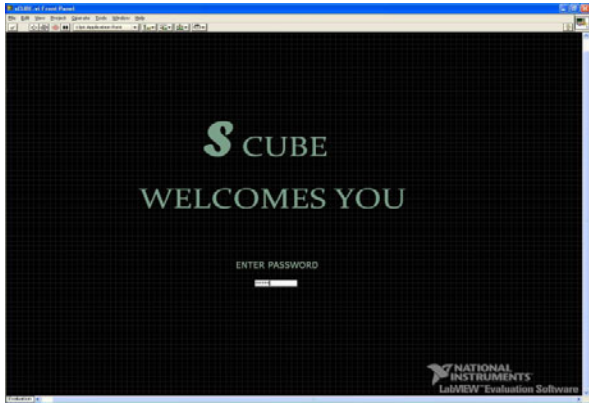


Fig. 10. Front panel (1) of the proposed system

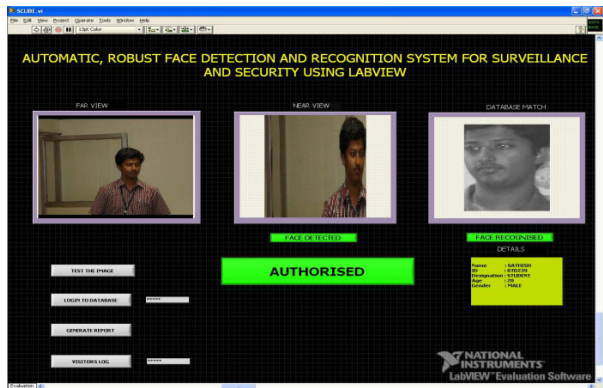


Fig. 11. Front panel (2) of the proposed system

7 Conclusion

We have thus successfully developed the system and tested the its working in the Biomedical Engineering Department at PSG College of Technology. We were able to achieve considerably high performance rate in surveillance using our proposed face detection and recognition methods. The successful implementation of the above automatic, robust face detection and recognition system using LabVIEW proves that with cutting edge technology and powerful software suites available nowadays, it is possible to implement any complex task, provided that the idea and the proposed methodology behind the task are defined very clearly.

References

1. Kanade, T.: Picture Processing by computer complex and recognition of human faces, PhD Thesis, Kyoto University (1973)
2. Brunelli, R., Poggio: Face Recognition: Features versus Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15(10), 1042–1052 (1993)
3. Turk, M.A., Pentland, A.P.: Face Recognition Using Eigenfaces. In: *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 586–591 (1991)
4. Demers, D., Cotrell, G.W.: Non-Dimensionality reduction. In: *Advances in Neural Information Processing Systems*, vol. 5, pp. 580–587. Morgan Kaufman Publishers, San Mateo (1993)
5. Turk, M., Pentland, A.: Eigenfaces for Recognition. *Journal of Cognitive Neuroscience* (March 1991)
6. Turk, M.A., Pentland, A.P.: Face recognition using eigenfaces. In: *Proc. of Computer Vision and Pattern Recognition*, pp. 586–591. IEEE, Los Alamitos (1991b)
7. Moon, H., Phillips, P.J.: Computational and Performance aspects of PCA-based Face Recognition Algorithms. *Perception* 30, 303–321 (2001)

Global Chaos Synchronization of Hyperchaotic Lorenz Systems by Sliding Mode Control

Sundarapandian Vaidyanathan¹ and Sivaperumal Sampath²

¹ R & D Centre, Vel Tech Dr. RR & Dr. SR Technical University
Avadi-Alamathi Road, Avadi, Chennai-600 062, India
sundarvtu@gmail.com

<http://www.vel-tech.org/>

² School of Electronics and Electrical Engineering, Singhania University
Dist. Jhunjhunu, Rajasthan-333 515, India
sivaperumals@gmail.com

<http://www.singhaniauniversity.co.in/>

Abstract. In this paper, new results based on the sliding mode control are derived for the global chaos synchronization of identical hyperchaotic Lorenz systems (Jia, 2007). The stability results for the sliding mode control based synchronization schemes derived in this paper are established using Lyapunov stability theory. Since the Lyapunov exponents are not required for these calculations, the sliding mode control method is very effective and convenient to achieve global chaos synchronization of the identical hyperchaotic Lorenz systems. Numerical simulations are shown to illustrate and validate the sliding mode control results derived in this paper for the identical hyperchaotic Lorenz systems.

Keywords: Sliding mode control, global chaos synchronization, chaos, hyperchaotic Lorenz system.

1 Introduction

Chaotic systems are dynamical systems that are highly sensitive to initial conditions. This sensitivity is popularly known as the *butterfly effect* [1]. Since the pioneering work by Pecora and Carroll ([2], 1990), chaos synchronization problem has been studied extensively in the literature. Chaos theory has been applied to a variety of fields including physical systems [3], chemical systems [4], ecological systems [5], secure communications ([6]-[8]) etc.

In the last two decades, various control schemes have been developed and successfully applied for the chaos synchronization such as PC method [2], OGY method [9], active control ([10]-[12]), adaptive control ([13]-[15]), time-delay feedback method [16], backstepping design method ([17]-[18]), sampled-data feedback synchronization method ([19]-[20]) etc.

In most of the chaos synchronization approaches, the *master-slave* or *drive-response* formalism is used. If a particular chaotic system is called the *master* or *drive* system and another chaotic system is called the *slave* or *response* system, then the goal of the global chaos synchronization is to use the output of the master system to control

the slave system so that the states of the slave system track the states of the master system asymptotically. In other words, global chaos synchronization is achieved when the difference of the states of master and slave systems converge to zero asymptotically with time.

In this paper, we derive new results based on the sliding mode control ([21]-[23]) for the global chaos synchronization of identical hyperchaotic Lorenz systems ([24], Jia, 2007). The stability results for the sliding mode control based synchronization schemes derived in this paper are established using Lyapunov stability theory [25]. In robust control systems, sliding mode control is often adopted due to its inherent advantages of easy realization, fast response and good transient performance as well as its insensitivity to parameter uncertainties and external disturbances.

This paper has been organized as follows. In Section 2, we describe the problem statement and our methodology using sliding mode control. In Section 3, we discuss the global chaos synchronization of identical hyperchaotic Lorenz systems ([24], Jia, 2007). In Section 4, we summarize the main results obtained in this paper.

2 Problem Statement and Our Methodology Using Sliding Mode Control

In this section, we detail the problem statement for global chaos synchronization of identical chaos systems and our methodology using sliding mode control (SMC) and Lyapunov stability theory.

Consider the chaotic system described by

$$\dot{x} = Ax + f(x) \quad (1)$$

where $x \in \mathbb{R}^n$ is the state of the system, A is the $n \times n$ matrix of the system parameters and $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is the nonlinear part of the system. We consider the system (1) as the *master* or *drive* system.

As the *slave* or *response* system, we consider the following chaotic system described by the dynamics

$$\dot{y} = Ay + f(y) + u \quad (2)$$

where $y \in \mathbb{R}^n$ is the state of the system and $u \in \mathbb{R}^m$ is the controller of the slave system.

If we define the *synchronization error* e as

$$e = y - x, \quad (3)$$

then the error dynamics is obtained as

$$\dot{e} = Ae + \eta(x, y) + u, \quad \text{where } \eta(x, y) = f(y) - f(x) \quad (4)$$

The objective of the global chaos synchronization problem is to find a controller u such that

$$\lim_{t \rightarrow \infty} \|e(t)\| = 0 \quad \text{for all initial conditions } e(0) \in \mathbb{R}^n \quad (5)$$

To solve this problem, we first define the control u as

$$u(t) = -\eta(x, y) + Bv(t) \tag{6}$$

where B is a constant gain vector selected such that (A, B) is controllable.

Substituting (6) into (4), the error dynamics simplifies to

$$\dot{e} = Ae + Bv \tag{7}$$

which is a linear time-invariant control system with single input v .

Thus, the original global chaos synchronization problem can be replaced by an equivalent problem of stabilizing the zero solution $e = 0$ of the linear system (7) by a suitable choice of the sliding mode control.

In the sliding mode control, we define the variable

$$s(e) = Ce = c_1e_1 + c_2e_2 + \dots + c_n e_n \tag{8}$$

where $C = [c_1 \ c_2 \ \dots \ c_n]$ is a constant vector to be determined.

In the sliding mode control, we constrain the motion of the system (7) to the sliding manifold defined by

$$S = \{x \in \mathbb{R}^n \mid s(e) = 0\} = \{x \in \mathbb{R}^n \mid c_1e_1 + c_2e_2 + \dots + c_n e_n = 0\}$$

which is required to be invariant under the flow of the error dynamics (7).

When in sliding manifold S , the system (7) satisfies the following conditions:

$$s(e) = 0 \tag{9}$$

which is the defining equation for the manifold S and

$$\dot{s}(e) = 0 \tag{10}$$

which is the necessary condition for the state trajectory $e(t)$ of the system (7) to stay on the sliding manifold S .

Using (7) and (8), the equation (10) can be rewritten as

$$\dot{s}(e) = C[Ae + Bv] = 0 \tag{11}$$

Solving (11), we obtain the equivalent control law given by

$$v_{eq}(t) = -(CB)^{-1}CAe(t) \tag{12}$$

where C is chosen such that $CB \neq 0$.

Substituting (12) into the error dynamics (7), we get the closed-loop dynamics as

$$\dot{e} = [I - B(CB)^{-1}C]Ae \tag{13}$$

where C is chosen such that the system matrix $[I - B(CB)^{-1}C]A$ is Hurwitz.

Then the controlled system (13) is globally asymptotically stable.

To design the sliding mode controller for the linear time-invariant system (7), we use the constant plus proportional rate reaching law

$$\dot{s} = -q\text{sgn}(s) - ks \tag{14}$$

where $\text{sgn}(\cdot)$ denotes the sign function and the gains $q > 0, k > 0$ are determined such that the sliding condition is satisfied and sliding motion will occur.

From equations (11) and (14), we obtain the control $v(t)$ as

$$v(t) = -(CB)^{-1}[C(kI + A)e + q\text{sgn}(s)] \tag{15}$$

Theorem 1. *The master system (1) and the slave system (2) are globally and asymptotically synchronized for all initial conditions $x(0), y(0) \in \mathbb{R}^n$ by the feedback control law*

$$u(t) = -\eta(x, y) + Bv(t) \tag{16}$$

where $v(t)$ is defined by (15) and B is a column vector such that (A, B) is controllable. Also, the sliding mode gains k, q are positive.

Proof. First, we note that substituting (16) and (15) into the error dynamics (7), we obtain the closed-loop dynamics as

$$\dot{e} = Ae - B(CB)^{-1}[C(kI + A)e + q\text{sgn}(s)] \tag{17}$$

To prove that the error dynamics (17) is globally asymptotically stable, we consider the candidate Lyapunov function defined by the equation

$$V(e) = \frac{1}{2} s^2(e) \tag{18}$$

which is a positive definite function on \mathbb{R}^n .

Differentiating V along the trajectories of (17) or the equivalent dynamics (14), we obtain

$$\dot{V}(e) = s(e)\dot{s}(e) = -ks^2 - q\text{sgn}(s) \tag{19}$$

which is a negative definite function on \mathbb{R}^n .

Thus, by Lyapunov stability theory [25], it is immediate that the error dynamics (17) is globally asymptotically stable for all initial conditions $e(0) \in \mathbb{R}^n$.

This completes the proof. ■

3 Global Chaos Synchronization of Identical Hyperchaotic Lorenz Systems

3.1 Main Results

In this section, we apply the sliding mode control results obtained in Section 2 for the global chaos synchronization of identical hyperchaotic Lorenz systems ([24], 2007).

Thus, the master system is described by the hyperchaotic Lorenz dynamics

$$\begin{aligned}
 \dot{x}_1 &= a(x_2 - x_1) + x_4 \\
 \dot{x}_2 &= -x_1x_3 + bx_1 - x_2 \\
 \dot{x}_3 &= x_1x_2 - cx_3 \\
 \dot{x}_4 &= -x_1x_3 + dx_4
 \end{aligned}
 \tag{20}$$

where x_1, x_2, x_3, x_4 are the states of the system and a, b, r, d are constant, positive parameters of the system.

The slave system is also described by the controlled hyperchaotic Lorenz dynamics

$$\begin{aligned}
 \dot{y}_1 &= a(y_2 - y_1) + y_4 + u_1 \\
 \dot{y}_2 &= -y_1y_3 + by_1 - y_2 + u_2 \\
 \dot{y}_3 &= y_1y_2 - cy_3 + u_3 \\
 \dot{y}_4 &= -y_1y_3 + dy_4 + u_4
 \end{aligned}
 \tag{21}$$

where y_1, y_2, y_3, y_4 are the states of the system and u_1, u_2, u_3, u_4 are the controllers to be designed.

The systems (20) and (21) are hyperchaotic when

$$a = 10, \quad b = 28, \quad c = 8/3 \quad \text{and} \quad d = 1.3$$

In this case, the system (20) has two positive Lyapunov exponents and hence, the system exhibits complex hyperchaotic behaviour. The state portrait of the hyperchaotic Lorenz system (20) is illustrated in Figure 1.

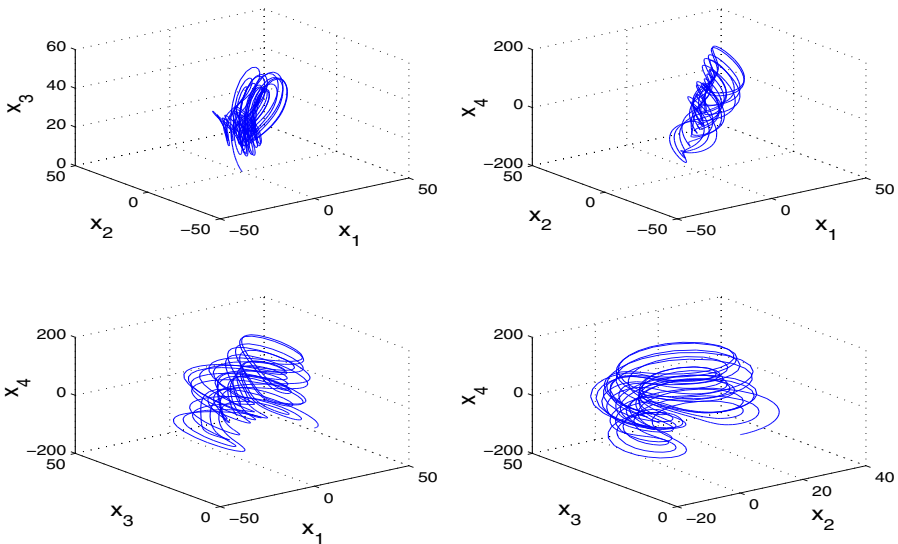


Fig. 1. State Portrait of the Hyperchaotic Lorenz System

The chaos synchronization error e is defined by

$$e_i = y_i - x_i, \quad (i = 1, 2, 3, 4) \quad (22)$$

The error dynamics is easily obtained as

$$\begin{aligned} \dot{e}_1 &= a(e_2 - e_1) + e_4 + u_1 \\ \dot{e}_2 &= be_1 - e_2 - y_1y_3 + x_1x_3 + u_2 \\ \dot{e}_3 &= -ce_3 + y_1y_2 - x_1x_2 + u_3 \\ \dot{e}_4 &= de_4 - y_1y_3 + x_1x_3 + u_4 \end{aligned} \quad (23)$$

We can write the error dynamics (23) in the matrix notation as

$$\dot{e} = Ae + \eta(x, y) + u \quad (24)$$

where the associated matrices are

$$A = \begin{bmatrix} -a & a & 0 & 1 \\ b & -1 & 0 & 0 \\ 0 & 0 & -c & 0 \\ 0 & 0 & 0 & d \end{bmatrix}, \quad \eta(x, y) = \begin{bmatrix} 0 \\ -y_1y_3 + x_1x_3 \\ y_1y_2 - x_1x_2 \\ -y_1y_3 + x_1x_3 \end{bmatrix} \quad \text{and} \quad u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} \quad (25)$$

The sliding mode controller design is carried out as detailed in Section 2.

First, we set u as

$$u = -\eta(x, y) + Bv \quad (26)$$

where B is chosen such that (A, B) is controllable. We take B as

$$B = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \quad (27)$$

In the hyperchaotic case, the parameter values are

$$a = 10, \quad b = 28, \quad c = 8/3 \quad \text{and} \quad d = 1.3$$

The sliding mode variable is selected as

$$s = Ce = [-1 \quad -2 \quad 0 \quad 1]e \quad (28)$$

which makes the sliding mode state equation asymptotically stable.

We choose the sliding mode gains as $k = 5$ and $q = 0.1$. We note that a large value of k can cause chattering and an appropriate value of q is chosen to speed up the time taken to reach the sliding manifold as well as to reduce the system chattering.

From equation (15), we can obtain $v(t)$ as

$$v(t) = -25.5e_1 - 9e_2 + 2.65e_4 + 0.05 \operatorname{sgn}(s) \quad (29)$$

Thus, the required sliding mode controller is obtained as

$$u(t) = -\eta(x, y) + Bv(t) \quad (30)$$

where $\eta(x, y)$, B and $v(t)$ are defined in equations (25), (27) and (29).

By Theorem 1, we obtain the following result.

Theorem 2. *The identical hyperchaotic Lorenz systems (20) and (21) are globally and asymptotically synchronized for all initial conditions with the sliding mode controller u defined by (30).* ■

3.2 Numerical Results

For the numerical simulations, the fourth-order Runge-Kutta method with time-step $h = 10^{-6}$ is used to solve the hyperchaotic Lorenz systems (20) and (21) with the sliding mode controller u given by (30) using MATLAB.

For the hyperchaotic Lorenz systems, the parameter values are taken as $a = 10$, $b = 28$, $c = 8/3$ and $d = 1.3$. The sliding mode gains are chosen as $k = 5$ and $q = 0.1$.

The initial values of the master system (20) are taken as

$$x_1(0) = 28, \quad x_2(0) = 12, \quad x_3(0) = 10, \quad x_4(0) = 20$$

and the initial values of the slave system (21) are taken as

$$y_1(0) = 15, \quad y_2(0) = 36, \quad y_3(0) = 34, \quad y_4(0) = 10$$

Figure 2 depicts the synchronization of the hyperchaotic Lorenz systems (20) and (21).

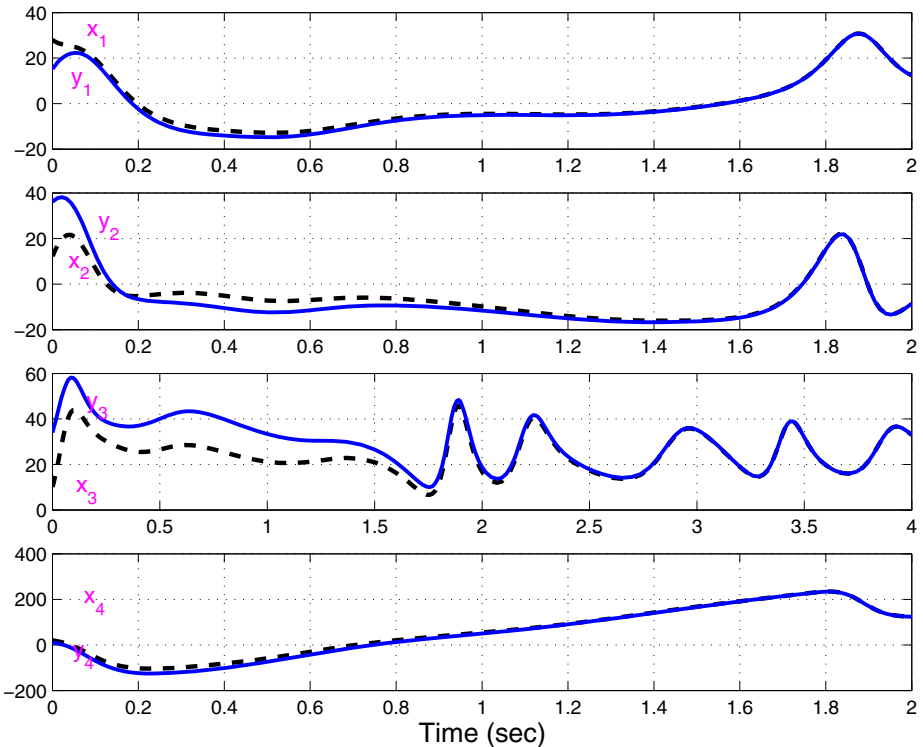


Fig. 2. Synchronization of the Identical Hyperchaotic Lorenz Systems

4 Conclusions

In this paper, we have used sliding mode control (SMC) to achieve global chaos synchronization for the identical hyperchaotic Lorenz systems (Jia, 2007). Our synchronization results for the identical hyperchaotic Lorenz systems have been established using Lyapunov stability theory. Since the Lyapunov exponents are not required for these calculations, the sliding mode control method is very effective and convenient to achieve global chaos synchronization for identical hyperchaotic Lorenz systems. Numerical simulations are also shown to illustrate the effectiveness of the synchronization results derived in this paper using sliding mode control.

References

1. Alligood, K.T., Sauer, T., Yorke, J.A.: *Chaos: An Introduction to Dynamical Systems*. Springer, New York (1997)
2. Pecora, L.M., Carroll, T.L.: Synchronization in chaotic systems. *Phys. Rev. Lett.* 64, 821–824 (1990)
3. Lakshmanan, M., Murali, K.: *Chaos in Nonlinear Oscillators: Controlling and Synchronization*. World Scientific, Singapore (1996)
4. Han, S.K., Kerrer, C., Kuramoto, Y.: Dephasing and bursting in coupled neural oscillators. *Phys. Rev. Lett.* 75, 3190–3193 (1995)
5. Blasius, B., Huppert, A., Stone, L.: Complex dynamics and phase synchronization in spatially extended ecological system. *Nature* 399, 354–359 (1999)
6. Kwok, H.S., Wallace, K., Tang, S., Man, K.F.: Online secure communication system using chaotic map. *Internat. J. Bifurcat. Chaos* 14, 285–292 (2004)
7. Kocarev, L., Parlitz, U.: General approach for chaos synchronization with applications to communications. *Phys. Rev. Lett.* 74, 5028–5030 (1995)
8. Murali, K., Lakshmanan, M.: Secure communication using a compound signal using sampled-data feedback. *Applied Math. Mech.* 11, 1309–1315 (2003)
9. Ott, E., Grebogi, C., Yorke, J.A.: Controlling chaos. *Phys. Rev. Lett.* 64, 1196–1199 (1990)
10. Ho, M.C., Hung, Y.C.: Synchronization of two different chaotic systems using generalized active network. *Phys. Lett. A* 301, 421–428 (2002)
11. Huang, L., Feng, R., Wang, M.: Synchronization of chaotic systems via nonlinear control. *Phys. Lett. A* 320, 271–275 (2004)
12. Chen, H.K.: Global chaos synchronization of new chaotic systems via nonlinear control. *Chaos, Solit. Fract.* 23, 1245–1251 (2005)
13. Chen, S.H., Lü, J.: Synchronization of an uncertain unified system via adaptive control. *Chaos, Solit. Fract.* 14, 643–647 (2002)
14. Lu, J., Han, X., Lü, J.: Adaptive feedback synchronization of a unified chaotic system. *Phys. Lett. A* 329, 327–333 (2004)
15. Samuel, B.: Adaptive synchronization between two different chaotic dynamical systems. *Adaptive Commun. Nonlinear Sci. Num. Simul.* 12, 976–985 (2007)
16. Park, J.H., Kwon, O.M.: A novel criterion for delayed feedback control of time-delay chaotic systems. *Chaos, Solit. Fract.* 17, 709–716 (2003)
17. Wu, X., Lü, J.: Parameter identification and backstepping control of uncertain Lü system. *Chaos, Solit. Fract.* 18, 721–729 (2003)
18. Yu, Y.G., Zhang, S.C.: Adaptive backstepping synchronization of uncertain chaotic systems. *Chaos, Solit. Fract.* 27, 1369–1375 (2006)

19. Yang, T., Chua, L.O.: Control of chaos using sampled-data feedback control. *Internat. J. Bifurcat. Chaos* 9, 215–219 (1999)
20. Zhao, J., Lu, J.: Using sampled-data feedback control and linear feedback synchronization in a new hyperchaotic system. *Chaos, Solit. Fract.* 35, 376–382 (2008)
21. Slotine, J.E., Sastry, S.S.: Tracking control of nonlinear systems using sliding surface with application to robotic manipulators. *Internat. J. Control* 38, 465–492 (1983)
22. Utkin, V.I.: Sliding mode control design principles and applications to electric drives. *IEEE Trans. Industrial Electr.* 40, 23–36 (1993)
23. Saravanakumar, R., Vinoth Kumar, K., Ray, K.K.: Sliding mode control of induction motor using simulation approach. *Internat. J. Control of Computer Science and Network Security* 9, 93–104 (2009)
24. Jia, Q.: Hyperchaos generated from the Lorenz chaotic system and its control. *Phys. Lett. A* 366, 217–222 (2007)
25. Hahn, W.: *The Stability of Motion*. Springer, New York (1967)

Segmentation of Micro Calcification Clusters in Digital Mammograms Using UIQI

Patnala S.R. Chandra Murty¹, T. Sudheer², and E. Sreenivasa Reddy³

^{1,2} Research Scholars

³ Professor

Department of Computer Science & Engg.,
Acharya Nagarjuna University, Guntur, A.P., India
Chandra_psr@rediffmail.com
edara_67@yahoo.com

Abstract. This paper presents a new method for automatic detection of clustered micro calcifications (both malignant and benign) in digitized mammograms. Compared to previous works, the innovation here is that the processing is performed by labeling the texture regions in the original image. This new method uses the labeling the textures of the mammographic images on the basis of UIQI (universal quality index). Once labeling is performed, mammogram is segmented into various regions and further thresholds are used to separate the clustered micro calcifications. By comparing our results with those found in the literature, we proved that the method enormously reduces the computing time as UIQI utilizes simple statistical features like mean, variance and standard deviations only. In this paper, an automatic segmentation and classification of massive lesions in mammographic images is evaluated on Mini-MIAS database consisting 320 digitalized mammograms. Furthermore, classification rates enhancements were also revealed by testing our method compared to existing methods.

Keywords: malignant, benign micro calcification, mammogram, UIQI.

1 Introduction

Breast cancer represents the most frequently diagnosed cancer in women. In order to reduce mortality, early detection of breast cancer is important, because diagnosis is more likely to be successful in the early stages of the disease. Women in the age group of 40-65 have more risk of breast cancer. In cold countries about 50%-90% of the women have this disease. In a Philippine study [1] mammogram screening was done to 151,198 women. Out of that 3479 women had this disease and were referred for diagnosis. In the European Union and the United States, it is the leading cause of death for women in their 40's [2] and second in Canada after lung cancer. In 2009, an estimated 22,700 Canadian women and 170 men were diagnosed with breast cancer and 5,400 women and 50 men died from it. Therefore, 1 in 9 women (11%) is expected to develop breast cancer during her lifetime (by age 90) and 1 in 28 will die from it. Mammography is the most effective procedure for an early diagnosis of the breast

cancer. Nowadays, people are trying to find a way or method to support as much as possible to the radiologists in diagnosis process. The most popular way is now being developed is using Computer-Aided Detection (CAD) system to process the digital mammograms and prompt the suspicious region to radiologist.

Micro calcifications (MC) are quiet minute bits of calcium, and may show up in clusters and are associated with extra cell activity in breast [5]. Usually the extra cell growth is not cancerous, but sometimes tight clusters of micro calcification can indicate early breast cancer. Those clustered micro calcifications are an important indicator for early detection of breast cancer [3]. Scattered micro calcifications are usually a sign of benign breast cancer. 80% of the MC is benign. MC in the breast shows up as white speckles on breast X-rays. Features extracted from mammograms can be used for detection of cancers [4]. Studies reports that features are extracted from the individual MCs [5] or from the ROI which contain MC clusters [6].

Joaquim.C. Felipe et al. [7] uses a set of shape based features. The paper presents the task of calcification and similarity retrieval of mammographic masses based on shape content. It also uses the statistical based association rule to discriminate the disease from the normal breast tissue. In [8] Chen et al. presents a new texture shape feature coding (TSFS) based classification method for classifying masses on mammograms. A texture shape histogram is used for generating various shape features of masses. Khuzi.A.M et al. [9] used a gray level co-occurrence matrix to provide the texture content information of Region of interest at different angles. Pelin Gorgel et al. [10] designed a wavelet based Support Vector Machine (SVM) for capturing information of the MCs. Decision making is done by extracting features as a first stage by computing wavelet coefficients and classification using the classifier trained on the extracted features. Prabhu shetty.K et al. [11] uses the spatial decomposition property of the Discrete Wavelet Transform and based on the statistical analysis the MC clusters are classified.

2 Universal Image Quality Index (UIQI)

The universal image quality index (UIQI) proposed by Wang and Bovik [20] measures how much distortion between original signal and reference image by modeling distortion as combination of three main components: correlation distortion, illumination distortion and contrast distortion.

The UIQI is a value that is obtained by comparing an image with the referenced one. Let..

$X = \{x_i \mid i = 1, 2, \dots, N.\}$ and $Y = \{y_i \mid i = 1, 2, \dots, N\}$ be the original and the test images respectively.

In order to construct a formula for UIQI, we require the values of following terms,

- a. Mean of image 'X'
- b. Mean of image 'Y'
- c. Variance of image 'X'
- d. Variance of image 'Y' and
- e. Standard Deviation of both 'X' and 'Y' images

UIQI is defined as:

$$\text{UIQI} = \frac{4 \sigma_{xy} \bar{X} \bar{Y}}{(\sigma_x^2 + \sigma_y^2)[(\bar{x})^2 + (\bar{y})^2]}$$

Where, mean of image X and image Y are

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N X_i, \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N Y_i$$

and the variance is calculated as

$$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2, \quad \sigma_y^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \bar{y})^2$$

Similarly standard deviation is

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$$

UIQI is the product of three quality measures reflecting these components, respectively: loss of correlation, luminance distortion, and contrast distortion.

3 Proposed Method

The proposed system uses UIQI based segment classification. The image considered from the mini MIAS database is first divided into segments of size 16 x 16, 8 x 8, 4x4 and so on. Out of these a segment of size 2 x 2 provides detailed edge information of the tumor. The segments are then used for UIQI calculation based on two adjacent segments, traversing from top-left sub image to the bottom-right sub image of original image. i.e. in horizontal direction.

Based on the resulting UIQI values, segments of the image can be clustered with pre specified upper and lower limits of the clusters, these clusters are then processed to color the segment containing the tumor with white and other segments with black. The segment with tumor part can be identified using pre-calculated UIQI values i.e., the segment with UIQI value near to 1 can be clustered and can be identified as tumor segments.

4 Experimental Results

Experimental evaluation of CAD system is carried on the mammogram images collected from mini MIAS Database. Mini MIAs Mammogram Data base contributes a total number of 320 mammogram images. Among them 32 images are used for our experiment out of which 25 are having one or more clusters of micro calcifications marked by expert radiologists. We have compared the effectiveness of proposed method with pixel coding and rank coding methods.

Fig 1 & 2 shows the segmented mammograms into sub images of size 2x2 and their rem with calcification regions segmented. Applying threshold values we have extracted both malignant and benign tumors as shown in Figure 6 & 7.

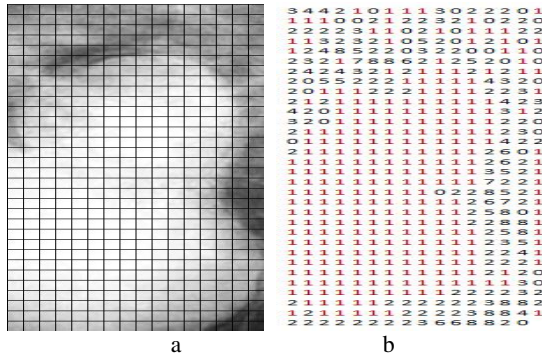


Fig. 1. Mammogram (a) Segmented into 3x3size (b) Labeled image by using UIQI

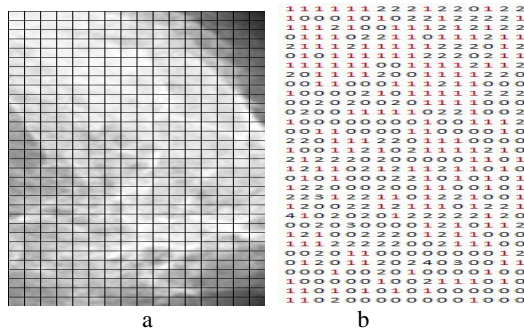


Fig. 2. Mammogram (a) Segmented into 3x3size (b) Labeled image by using UIQI

For Benign Tumor:

Threshold for UIQI in labeling:

```

if(uiqi[i]>-10 && uiqi[i]<=30) // upper and lower limits for UIQI
{
    lbl[i]=1; //class label is 1
}
    
```

Threshold for segmentation:

```

if(avg>140 && avg<200)
yimg.setRGB(j,i,RGB); //tumor pixel values are set as it is
else
yimg.setRGB(j,i,0xff000000); //other pixels are set to "0"
    
```

For Malignant Tumor:

Threshold for UIQI in labeling:

```

if(uiqi[i]>-10 && uiqi[i]<=30) // upper and lower limits for UIQI
{
    lbl[i]=1; //class label is 1
}
    
```

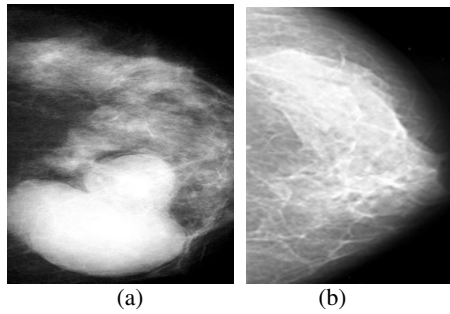



Fig. 3. a) Original mammogram Images b) Calcification regions identified

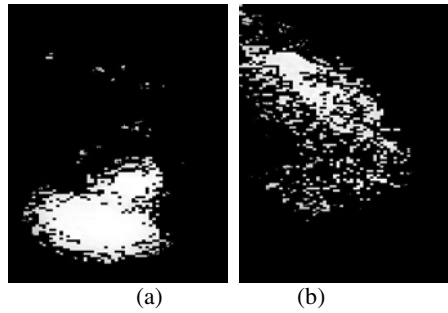


Fig. 4. a) Original mammogram Images b) Calcification regions identified

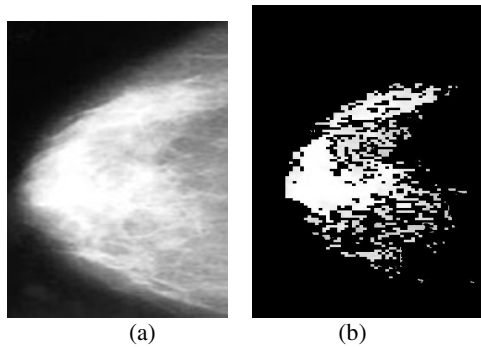


Fig. 5. a) Original mammogram Images b) Calcification regions identified

Threshold for segmentation:

```

if(avg>188)
yimg.setRGB(j,i,RGB); //tumor pixel values are set as it is
else
yimg.setRGB(j,i,0xff000000); //other pixels are set to "0

```

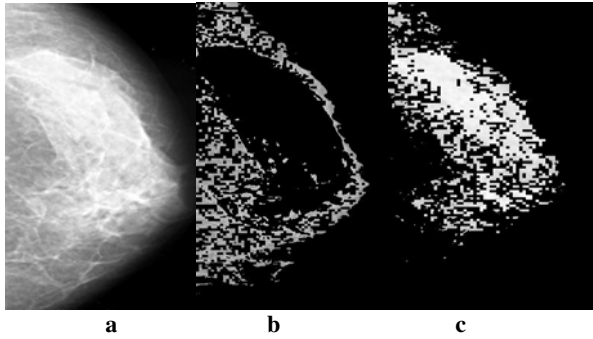


Fig. 6.

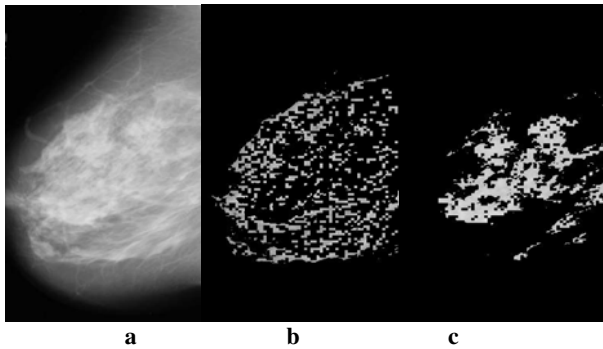


Fig. 7.

Table 1. Comparative study of Classification rates

Method	Test Cases	True +ve	True -ve	False +ve	False -ve
Using UIQI	No of Cases	24/25	6/7	1/25	1/7
	%age	96%	86%	4%	14%
Using Simple Coding	No of Cases	18/25	5/7	7/25	2/7
	%age	72%	71%	28%	39%
Using Rank Coding	No of Cases	20/25	6/7	5/25	1/7
	%age	80	86%	20%	14%

Both rank coding and texture coding methods considers only limited range of values. This reduction is use full in the view of space complexity, but it suffers more from segmentation point of view. In terms of classification, we consider statistical texture characteristics while UIQI. Table 1 summarizes the percentage of classified rates.

The classification rates obtained confirm that superior performance is obtained by existing method. As we can change window size (sub image) it has variable computational complexity. We can select a suitable window size for a given classification rate to minimize complexity.

5 Conclusion

We have developed a new method for detection of both benign and malignant micro calcifications based on labeling and thresholding. Each image is subdivided as small as 2 x2 sub images and their UIQI is found. Based on UIQI index labeling is performed for different regions in the mammogram. Applying proper empirical threshold values we can extract benign calcifications. Experimental results showed that calcification rates found by our method is 96% where as that of using texture coded images is 92% and that of using rank codes is around 80%. As the segmentation is done at the level of sub image, the method is easier for implementation and reduces computational complexity when compared with pixel coding and gray level coding methods.

References

- [1] Pisani, et al.: Outcome of screening by Clinical Examination of the Breast in a Trial in the Phillipines. *Int. J. Cancer* (2006)
- [2] Oliver, A., Freixenet, J., Marti, R., Pont, J., Perez, E., Denton, E.R.E., Zwiggelaar, R.: A novel breast tissue density classification methodology. *IEEE Transactions on Information Technology in Biomedicine* 12(1), 55–65 (2008)
- [3] Eddaoudi, F., Regragui, F., Laraki, K.: Characterization of normal mammograms based on the Statistical Features. In: *Proc. Second International Symposium on Communications, Control and Signal Processing (ISCCSP), Marrakech* (2006)
- [4] Shen, L., Rangaan, R.M., Desautels, J.E.L.: Application of shape analysis to mammographic classifications. *IEEE Trans. Med. Imag.* 13(2), 224–263 (1994)
- [5] Lee, S.K., Chung, P., Chang, C.L., Lo, C.S., Lee, T., Hsu, G.C., Ang, C.W.: Classification of Clustered Microcalcifications Using Shape Cognitron Neural Network. *Neural Netw.* 16(1), 121–132 (2003)
- [6] Dhawan, A.P., Chitre, Kaiser-Bonasso, C., Moskoitz, M.: Analysis of mammographic microcalcification using gray-level image structure features. *IEEE Trans. Med. Imag.* 15(3), 11–150 (2005)
- [7] Felipe, J.C., et al.: Effective shape based retrieval and classification of mammograms. In: *Proceeding of the ACM Smposium in Applied Computing* (2006)
- [8] Chen, Chang, C.: New Texture shape feature coding based computer aided diagnostic methods for classification of masses on mammograms. In: *Engg. In Medicine and Biology Society IEMBS 26th Annual Int. Conference of the IEEE, vol. 1, pp. 1275–1128* (2004)
- [9] Khuzi, A.M., Besar, R., an Zaki, M.D.: Texture feature selection for masses detection in digital mammograms. In: *IFMBE Proceedings Springerlink, vol. 21, pp. 629–632*
- [10] Gorgel, P., Serlbas, A., et al.: Mammogram mass classification using wavelet based support vector machine. *Journal of Electrical and Electronics Engg.* 9(1), 867–875 (2009)
- [11] Prabhu Shetty, K., Udupi, V.R.: Wavelet based microcalcification detection on mammographic images. *IJCSNS* 9(7) (July 2009)

- [12] Baraldi, Parmiggiani, F.: An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters. *IEEE Transactions on Geoscience and Remote Sensing* 33(2), 293–304 (1995)
- [13] Bastos, C., Ferreira, R., Borges, D.L.: Analysis of mammogram classification using a wavelet transform decomposition. *Pattern Recognition Letters* 24(7), 973–982 (2003)
- [14] Sumeet, D., Harpreet, S., Thompson, H.W.: Associative classification of mammograms using weighted rules. *Expert Systems with Applications* 36, 9250–9259 (2009)
- [15] Júlia, E.E.O., Deserno, T.M., Araújo, A.: Breast Lesions Classification applied to a reference database. In: 2nd International Conference: E-Medical Systems, Tunisia, October 29–31 (2008)
- [16] Eddaoudi, F., Regragui, F., Laraki, K.: Characterization of normal mammograms based on the Statistical Features. In: Proc. Second International Symposium on Communications, Control and Signal Processing (ISCCSP), Marrakech (2006)
- [17] Bentley, K., et al.: Mammography image quality: Analysis of evaluation criteria using pectoral muscle presentation. *Radiography* (2007), doi:10.1016/j.radi.2007.02.002
- [18] Heucke, L., Knaak, M., Orglmeister, R.: A new image segmentation method based on human brightness perception and foveal adaptation. *IEEE Signal Processing Letters* 7(6), 129–131 (2000)
- [19] Whi-Vin, O., KwangGi, K., Young-Jae, K., HanSung, K., JungSil, R., WooKyung, M.: Detection of Microcalcifications in Digital Mammograms Using Foveal Method. *Journal of Korean Society of Medical Informatics* 15-1, 165–172 (2009)
- [20] Gorgel, P., Sertbas, A., Kilic, N., Ucan, O.N., Osman, O.: Mammographic mass classification using wavelet based support vector machine. *Journal of Electrical & Electronics Engineering* 9(1), 867–875 (2009)
- [21] Bovik, A.C., Wang, Z.: A Universal Image Quality Index. *IEEE Signal Processing Letters* 9(3), 81–84 (2002)
- [22] Analysis of Relative Entropy, Accuracy, and Quality of Face Biometric. In: Appeared in the Pattern Recognition for IT Security Workshop, Darmstadt, Germany (September 2010)

A Denoising Approach for Salt and Pepper Noise Corrupted Image at Higher Noise Density

Debika Dey, Samanwita Laha, Shalini Chowdhury, and Subhojit Sarker

Department of Electronics & Communication Engineering,
Siliguri Institute Of Technology, Sukna, Darjeeling, India
{debika.dey1,shln.chowdhury,samanwita30,
subhojitsarker}@gmail.com
<http://www.gmail.com>

Abstract. Image restoration and noise reduction is an eminent problem in almost all image processing applications. Numerous image restoration methods have been developed, each of which has its own advantages and limitation. This paper proposes a novel approach for removal of salt and pepper noise using a two stage process, in which the noisy image is first subjected to an adaptive median filter and then its output is further denoised by applying it to a new patch based non- local recovery paradigm. The Non-Local means filter uses the redundancy of information in the image under study to remove the noise. The statistical results of simulations are done using MATLAB and the obtained denoised images are quantified using various performance metrics.

Keywords: Adaptive median Filter, Non-Local means filter, Salt and pepper noise, Image restoration.

1 Introduction

Salt and pepper noise is caused by defective pixel in camera sensors and often found in digital transmission and storage. For images corrupted by this noise, the noisy pixels can take only the maximum and the minimum values (respectively, any random value) in the dynamic range [1]. Salt and Pepper noise is a special type of impulse noise. The probability density function (PDF) is

$$P(z) = \begin{cases} P_a, & z = a \\ P_b, & z = b \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

If neither of the probability is zero then the impulse noise resembles salt and pepper granules, distributed randomly over the image, hence the name. The removal of salt and pepper noise is generally approached using median type filters [2]. Previously Standard Median Filtering technique used to be considered as a robust technique in terms of noise attenuation and edge preservation. However in this method, when the noise variance is more than 0.5, some details and edges of the image are smashed [3]. An appropriate method of salt and pepper reduction is one which increases signal to noise ratio while preserving the edges and other fine details. To achieve this, an

adaptive structure of the median filter was developed [4]. This adaptive median filter ensures that most of the impulse noise is detected even at a high noise level provided that the window size is large enough. This method too seized to yield low error at higher noise variance i.e., about 0.7 and blurring of image becomes prominent [5].

To obtain a lower MSE and higher SNR, this paper proposes a two phase filtering technique. The noisy image is first subjected to a standard Adaptive Median Filter [6]. The filtered image is then denoised using Non-Local Means Filtering technique [7]. The NLM filter was introduced by Buades in 2005 [8]. This method of image denoising relies on the weighted average of all pixel intensities where the family of weights depends on the similarity between the pixels and the neighborhood of the pixel being processed. The proposed method outperforms the Standard Median Filter as well as Adaptive Median Filter in terms of several performance parameters, at higher noise variances. The numerical result obtained supports this claim.

The rest of the paper is organized as follows. The background is presented in Section 2, which includes the techniques for Median filter, Adaptive Median Filter as well as the Non-Local Means Filtering. Then in Section 3, the proposed method is described and finally Section 4 and 5, reports the simulation results and concluding remarks respectively.

2 Background

2.1 Standard Median Filter

The Standard Median Filter considers each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. Instead of simply replacing the pixel value with the *mean* of neighboring pixel values, it replaces it with the *median* of those values. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value. (If the neighborhood under consideration contains an even number of pixels, the average of the two middle pixel values is used.) Figure. 1 illustrates an example calculation [9] [10].

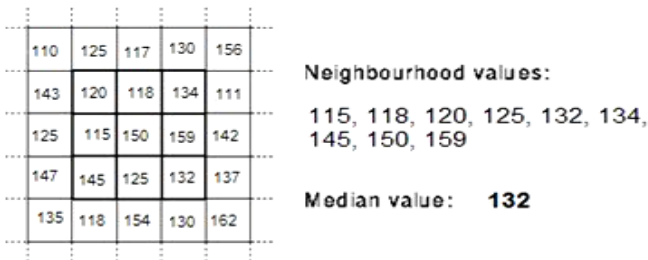


Fig. 1. Calculating the median value of a pixel neighborhood

As can be seen in figure 1, the central pixel value of 150 is rather unrepresentative of the surrounding pixels and is replaced with the median value: 132. A [3×3] square neighborhood is used here. Larger neighborhoods will produce more severe smoothing.

2.2 Adaptive Median Filter

The standard median filter performs well as long as the spatial noise density of the salt and pepper noise is not large. The filter performance degrades when the spatial noise variance of the salt and pepper noise increases [3]. Further with larger image and as the size of the kernel increases, the details and the edges becomes obscured [11]. The standard median filter does not take into account the variation of image charecteristics from one point to another. The behavior of adaptive filter changes based on statistical characteristic of the image inside the filter region defined by the $m \times n$ rectangular window S_{xy} [11]. The adaptive median filter differs from other adaptive filter as the size of the rectangular window S_{xy} is made to vary depending on

- (a) z_{min} =Minimum gray level value in S_{xy}
- (b) z_{max} =Maximum gray level value in S_{xy}
- (c) z_{med} =Medians of gray level in S_{xy}
- (d) z_{xy} =Gray levels at coordinate (x,y)
- (e) S_{max} = Maximum allowed size of S_{xy} [12].

The flowchart for adaptive median filtering is based on two levels is shown in figure 2.

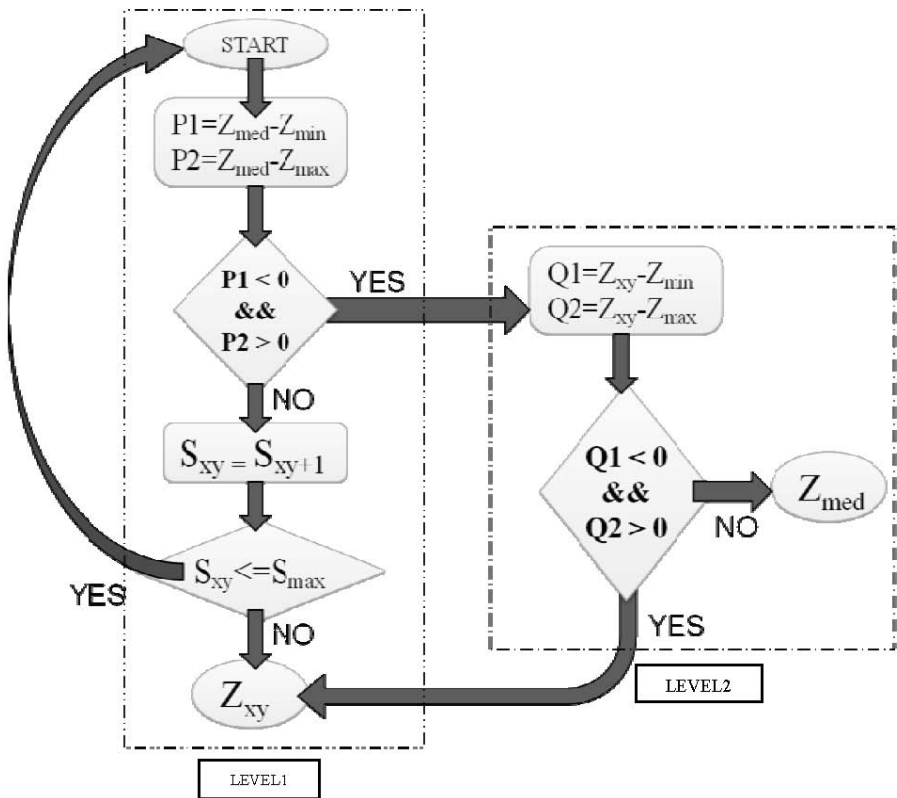


Fig. 2. Flowchart of Adaptive median filter

The adaptive median filtering algorithm works in two levels, denoted by LEVEL1 and LEVEL2. The use of AMF serves three main purpose: to remove salt and pepper (impulse) noise; to provide smoothing of other noise that may not be impulsive, and to reduce distortion, such as excessive thinning or thickening of object boundaries. The values Z_{\min} and Z_{\max} are considered statistically by the algorithm to be ‘impulse like’ noise components, even if these are not the lowest and highest possible pixel values in the image.

The purpose of LEVEL1 is to determine if the median filter output Z_{med} is impulse output or not. If LEVEL1 does find an impulse output then that would cause it to branch to LEVEL2. Here, the algorithm then increases the size of the window and repeats LEVEL1 and continues until it finds a median value that is not an impulse or the maximum window size is reached, the algorithm returns the value of Z_{xy} .

Every time the algorithm outputs a value, the window S_{xy} is moved to the next location in the image. The algorithm is then reinitialized and applied to the pixels in the new location. The median value can be updated iteratively using only the new pixels, thus reducing computational overhead.

2.3 Non-local Means Filtering

The approach of Non-local means filter was introduced by Buades in 2005 [7] based on non-local averaging of all pixels in the image. The method was based on denoising an image corrupted by white gaussian noise with zero mean and variance.

The Non Local Means filtering approach estimates every pixel intensity based on information from the whole image thereby exploiting the presence of similar patterns and features in the image. In this method, the restored gray value of each pixel is obtained by the weighted average of the gray values of all pixels in the image. Each weight is proportional to the similarity between the local neighborhood of the pixel being processed and the neighborhood corresponding to the other image pixel [8].

Given a discrete noisy image $v = \{v(i)\}$ for a pixel I the estimated value of $NL[v](i)$ is computed as weighted average of all the pixels i.e.:

$$NL[v](i) = \sum_{j \in i} w(i, j) \cdot v(j) \quad (2)$$

where the family of weights $\{w(i, j)\}$ depend on the similarity between the pixels i and j ,

The similarity between two pixels i and j depends on the similarity of the intensity gray level vectors $v(N_i)$ and $v(N_j)$, where N_k denotes a square neighborhood of fixed size and centered at a pixel k . The similarity is measured as a decreasing function of the weighted Euclidean distance, $\|v(N_i) - v(N_j)\|_{2, \alpha}^2$ where $\alpha > 0$ is the standard deviation of the Gaussian kernel.

The pixels with a similar grey level neighborhood to $v(N_i)$ have larger weights in the average.

These weights are defined as,

$$w(i, j) = \frac{1}{z(i)} e^{-\frac{\|v(N_i) - v(N_j)\|_{2, \alpha}^2}{h^2}} \quad (3)$$

where $Z(i)$ is the normalizing constant and the parameter h acts as a degree of filtering. It controls the decay of the exponential function and therefore the decay of the weights as a function of the Euclidean distances.

3 Proposed Method

The original test image is corrupted with simulated salt and pepper noise with different noise variance ranging from 0.1 to 0.9. In the proposed denoising approach, the noisy image is first applied to an adaptive median filter. The maximum allowed size of the window of the adaptive median filter is taken to be 5×5 for effective filtering. The choice of maximum allowed window size depends on the application but a reasonable value was computed by experimenting with various sizes of standard median filter.

In the second stage the resultant image is subjected to NL-means filtering technique.

The block diagram of the proposed method is shown in figure 3 below:

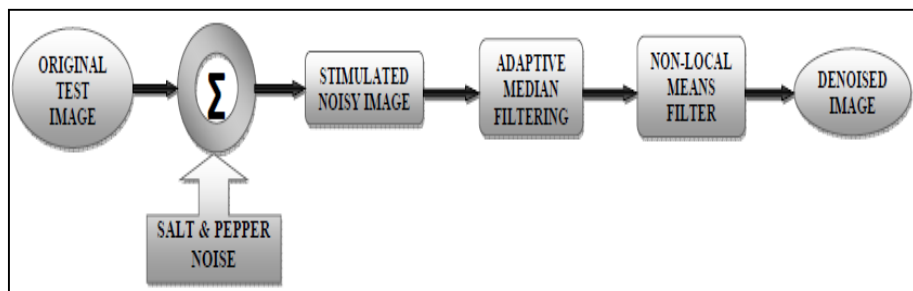


Fig. 3. Block diagram of proposed method

The performance of the filter depends on,

- radio of search window (value taken 7)
- radio of similarity window (value taken 5)

- degree of filtering (taken equal to the value of noise variance divided by 5).

The performance of the proposed technique is quantified by using various performance metrics such as, mean average error (MAE), mean square error (MSE), signal to mean square error (S/MSE), signal to noise ratio (SNR) and peak signal to noise ratio (PSNR).

4 Experimentation

4.1 Simulation

Intensive simulations were carried out using several monochrome images, from which a 256×256 "Lena.png" and "Cameraman.tif" image is chosen for demonstration. The test image is corrupted by fixed value salt and pepper noise with noise variance varying from 0.1 to 0.9.

In this section the output of the proposed technique is compared with different standard methods such as Median filters MF (3X3) and MF (5X5), Adaptive median filter AMF (7X7). The results are quantified using the following well defined parameters such as,

A) Mean Average Error (MAE) [13]

$$MAE = \frac{1}{MN} \sum_i \sum_j |r_{ij} - x_{ij}| \quad (4)$$

where r_{ij} and x_{ij} denote the pixel values of the restored image and the original image respectively and $M \times N$ is the size of the image.

B) Mean Square Error (MSE)

$$MSE = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [I(i,j) - \hat{I}(i,j)]^2 \quad (5)$$

Where M and N are the total number of pixels in the horizontal and the vertical dimensions of the image. I and \hat{I} denote the original and filtered image, respectively [14].

C) Signal to Mean Square Error(S/MSE) [15]

$$S/MSE = 10 \log_{10} \frac{\sum_{i=1}^K S_i^2}{\sum_{i=1}^K (\hat{S}_i - S_i)^2} \quad (6)$$

where \hat{S}_i and S_i denote the pixel values of the restored image and original image respectively.

D) Peak Signal to Noise Ratio [14]

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right) \quad (7)$$

E) Signal to Noise Ratio (SNR) [15]

$$SNR = 10 \log_{10} \frac{\sum_{x_i \in \Omega^2} (v(x_i)^2 - \hat{v}(x_i)^2)}{\sum_{x_i \in \Omega^2} (v(x_i) - \hat{v}(x_i))^2} \quad (8)$$

where $v(x_i)$ is the true value of pixel x_i and $\hat{v}(x_i)$ is the restored value of pixel x_i .

4.2 Denoising Performance

In this paper, salt and pepper noise was added to the test image with noise variance ranging from 0.1 to 0.9. The results for noisy and denoised image are shown in Table 1 for Lena image.

The performance of the proposed algorithm is tested for various levels of noise variance in Cameraman image and compared with standard filters namely Standard Median Filter (MF) of window size (3x3) and (5x5), Adaptive Median Filter (AMF), in terms of Mean Absolute Error (MAE) [13], Signal-To-Noise Ratio (SNR) [15], and the results are shown in Table 2 and 3 respectively .



Fig. 4. (A) Original 'Lena.png' image (B) Noisy image with variance-0.6 (C) Median Filter (3x3) output (D) Median Filter (5x5) output (E) Adaptive median filter output (F) Denoised image using proposed method

The analysis of the results obtained in Table 1 show that the Mean Absolute Error (MAE) and the Mean Square Error (MSE) is low for images corrupted with salt and pepper noise of low variance. However with the increase in noise variance there is a corresponding increase in calculated error. It can be seen that error in the denoised image has been reduced to a large extent as compared to the noisy image. There is an

Table 1. Performance matrices of the proposed method for Lena image

PERFORMANCE METRICS--- lena.png						
VARIANCE	IMAGE TYPE	MAE	MSE	SMSE	SNR	PSNR
0.1	NOISY	0.0495	0.0306	9.2031	11.9808	63.2751
	PROPOSED METHOD	0.0144	0.0007	25.3180	28.3409	79.8937
0.2	NOISY	0.1004	0.0621	6.5932	9.1570	60.1993
	PROPOSED METHOD	0.0181	0.0009	23.8311	26.8580	78.4148
0.3	NOISY	0.1506	0.0932	5.2216	7.6156	58.4372
	PROPOSED METHOD	0.0225	0.0014	22.0297	25.0610	76.6221
0.4	NOISY	0.2009	0.1245	4.3316	6.5736	57.1797
	PROPOSED METHOD	0.0269	0.0020	20.5609	23.6019	75.1725
0.5	NOISY	0.2514	0.1558	3.6870	5.7985	56.2038
	PROPOSED METHOD	0.0307	0.0026	19.3449	22.3949	73.9745
0.6	NOISY	0.3004	0.1860	3.2097	5.2117	55.4361
	PROPOSED METHOD	0.0352	0.0034	18.1689	21.2410	72.8421
0.7	NOISY	0.3498	0.2165	2.8486	4.7424	54.7759
	PROPOSED METHOD	0.0425	0.0055	16.0478	19.1321	70.7449
0.8	NOISY	0.4010	0.2482	2.5046	4.3122	54.1835
	PROPOSED METHOD	0.0633	0.0136	12.1550	15.2175	66.8092
0.9	NOISY	0.4482	0.2770	2.3315	4.0381	53.7066
	PROPOSED METHOD	0.1448	0.0522	7.0313	9.7375	60.9507

improvement in SNR and PSNR as well. For instance for noise variance of 0.7 there is an improvement of PSNR by about 26db.

From Table 2-3 it is evident that the Median Filter performs best at lower noise variance (upto 0.4) i.e. the SNR value is highest and also a low value of MAE has been achieved. But when noise variance is increased to about 0.6, the MF performance degrades and Adaptive Median Filter suppresses much of the noise but it

produces streaking effect and not suitable for noise variance of above 0.65. Experimental results show that the proposed method restores the original image much better than standard non-linear median-based filter and AMF at higher noise variance. For instance at noise variance of 0.7 the SNR of the restored image improves by about 14db as compared to the noisy image as opposed to the case of 11db for AMF.

The performance of the various filters with respect to MAE and SNR are plotted in Figure 5 and 6 respectively.

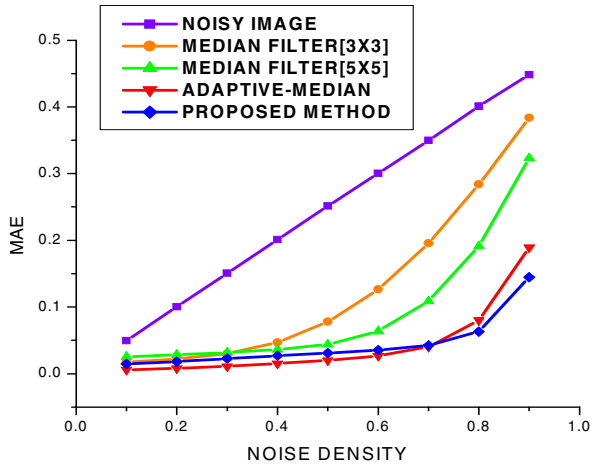


Fig. 5. Graph showing comparison of MAE obtained for different Noise variance of Lena image

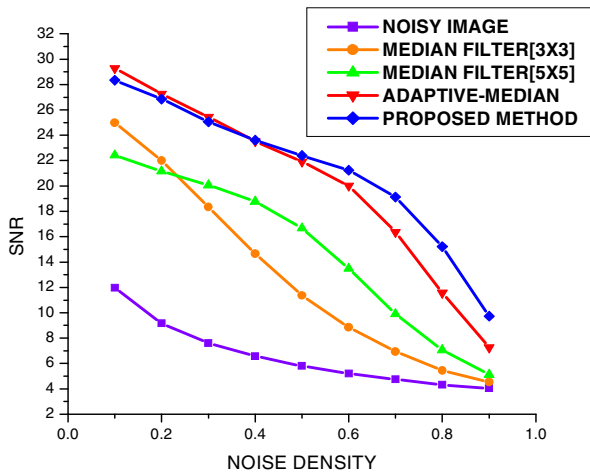


Fig. 6. Graph showing comparison of SNR obtained for different Noise variance of Lena image

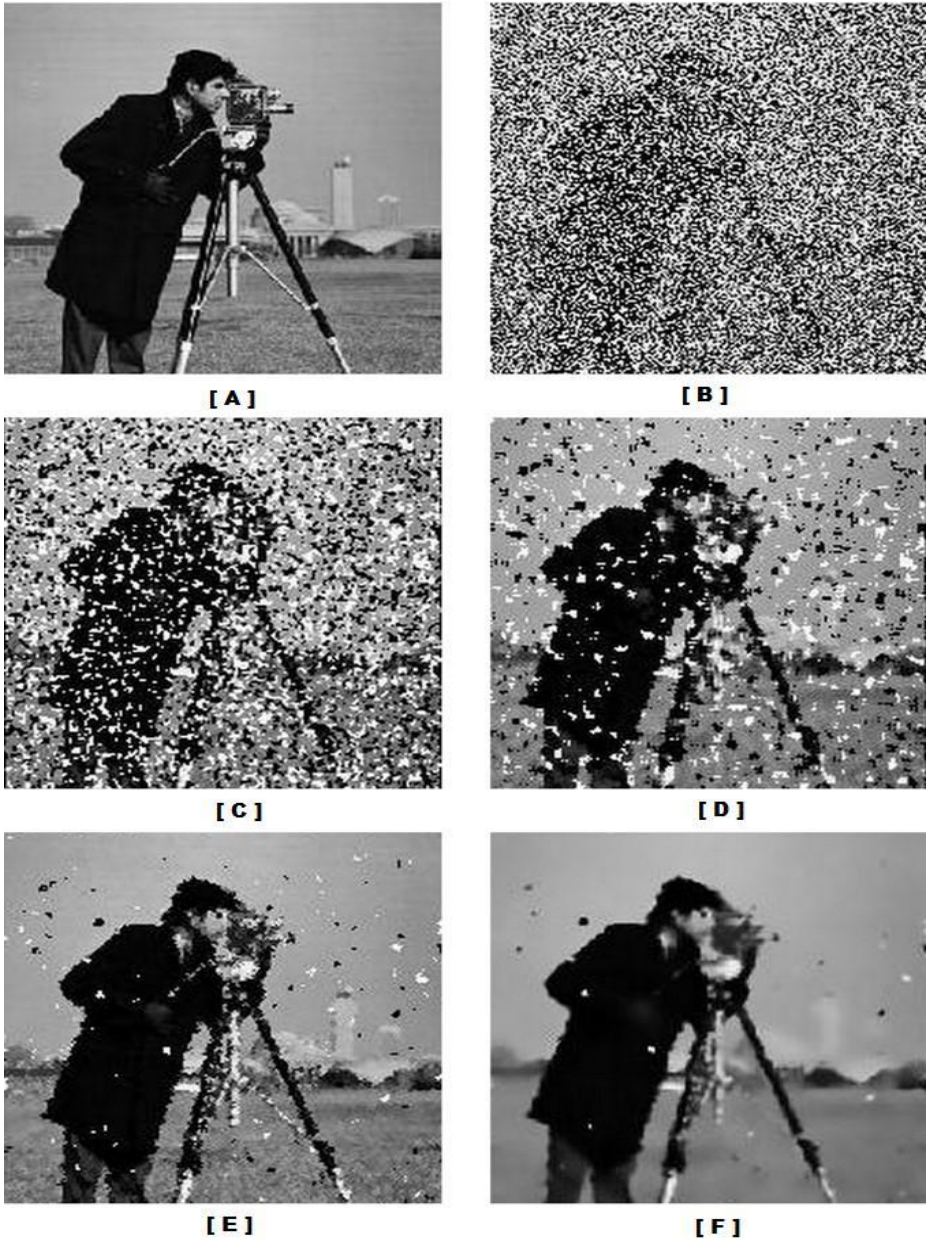


Fig. 7. (A) original Cameraman.tif image (B) Noisy image with variance-0.7 (C) Median Filter (3x3) output (D) Median Filter (5x5) output (E) Adaptive median filter output (F) Denoised image using proposed method

Table 2. Comparison of MAE values with other existing methods for a Cameraman.tif image

FILTER TYPE	MAE—cameraman.tif		
	$\sigma = 0.4$	$\sigma = 0.6$	$\sigma = 0.7$
NOISY	0.1992	0.2980	0.3472
MEDIAN-[3X3]	0.0486	0.1268	0.1942
MEDIAN-[5X5]	0.0387	0.0646	0.1061
ADAPTIVE-MEDIAN	0.0153	0.0271	0.0378
PROPOSED METHOD	0.0253	0.0343	0.0381

Table 3. Comparison of SNR values with other existing methods for a Cameraman.tif image

FILTER TYPE	SNR-- cameraman.tif		
	$\sigma = 0.4$	$\sigma = 0.6$	$\sigma = 0.7$
NOISY	7.0160	5.5565	5.0436
MEDIAN-[3X3]	14.6698	9.2742	7.3749
MEDIAN-[5X5]	17.3610	13.3020	10.3452
ADAPTIVE-MEDIAN	22.5668	18.8561	16.6243
PROPOSED METHOD	23.0670	20.1412	19.1356

5 Conclusion

In this paper a new method is developed for restoration of an image, corrupted by salt and pepper noise. The existing filters like median filter and adaptive median filter can denoise salt and pepper noise for lower values of noise variance, but fail to remove noise effectively if the noise variance increases. This paper proposes a method to handle salt and pepper noise having higher variances i.e. more than 0.65.

To demonstrate the performance of the proposed method, extensive simulation experiments have been conducted on a variety of standard test images to compare the proposed method with many other well known techniques. Experimental results simulated with MATLAB 7 indicate that the proposed method performs significantly better than many other existing techniques when the noise variance is higher and these results are also graphically analyzed for comparative study.

Although the proposed method outperforms the existing denoising techniques at higher value of noise variances, but scope of improvement still exists. As this method involves a two stage process, the number of computations is very high and hence the simulation time increases with increase in the size of the corrupted image. So as a future work, modifications may be incorporated to reduce the computation time and make the algorithm faster.

References

1. Chan, R.H., Ho, C.-W., Nikolova, M.: Salt-and-Pepper Noise Removal by Median-Type Noise Detectors and Detail-Preserving Regularization. *IEEE Transactions on Image Processing* 14(10) (October 2005)
2. Astola, J., Kuosmanen, P.: *Fundamentals of Nonlinear Digital Filtering*. CRC, Boca Raton (1997)
3. Chen, T., Whu, H.R.: Space Space variant median filters for the restoration of impulse noise corrupted images. *IEEE Trans. Image Processing* 7, 784–789 (1998)
4. Hwang, H., Haddad, R.A.: Adaptive median filters: new algorithms and results. *IEEE Transactions on Image Processing* 4, 499–502 (1995)
5. Jayaraj, V., Ebenezer, D., Aiswarya, K.: High Density Salt and Pepper Noise Removal in images using Improved Adaptive Statistics Estimation Filter. *IJCSNS International Journal of Computer Science and Network Security* 9(11) (November 2009)
6. Gonzalez, R.C., Woods, R.E., Eddins, S.L.: *Digital Image Processing Using MATLAB*. Prentice-Hall, Englewood Cliffs (2004)
7. Buades, A., Coll, B., Morel, J.M.: A review of image denoising algorithms with a new one. *Society for Industrial and Applied Mathematics* 4(2), 490–530 (2005)
8. Buades, A., Coll, B., Morel, J.M.: A non local algorithm for image denoising. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2005*, vol. 2, pp. 60–65 (2005)
9. Ko, S.-J., Lee, Y.H.: Center weighted median filters and their applications to image enhancement. *IEEE Trans. Circuits Syst.* 38(9), 984–993 (1991)
10. Sun, T., Neuvo, Y.: Detail-preserving median based filters in image processing. *Pattern Recognit. Lett.* 15(4), 341–347 (1994)
11. Maragos, P., Schafer, R.: Morphological Filters—Part II: Their Relations to Median, Order-Statistic, and Stack Filters. *IEEE Trans. Acoust., Speech, Signal Processing* 35(8), 1170–1184 (1987)
12. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*, 2nd edn. Prentice-Hall, Englewood Cliffs (2002)
13. Vijaykumar, V.R., Vanathi, P.T., Kanagasabapathy, P., Ebenezer, D.: Robust Statistics Based Algorithm to Remove Salt and Pepper Noise in Images. *International Journal of Information and Communication Engineering* 5, 3 (2009)
14. Juneja, M., Sandhu, P.S.: Design and Development of an Improved Adaptive Median Filtering Method for Impulse Noise Detection. *International Journal of Computer and Electrical Engineering* 1(5) (December 2009)
15. Sarker, S., Devi, S.: Effect of Non-Local Means filter in a Homomorphic Framework Cascaded with Bacterial Foraging Optimization to Eliminate Speckle. *CiiT International Journal of Digital Image Processing* 3(2) (February 2011)

Blind Medical Image Watermarking Technique for Secure Recovery of Hidden Data

Shaik. Basheera¹, D. Bhanu Prakash², and P.V. Naganjaneyulu³

¹ Eswar College of Engineering, Narasaraopet, Guntur (Dist), Andhra Pradesh, India
basher_405@rediffmail.com

² Sri Chaitanya Engineering College, Krimnagar, Karimnagar (Dist), Andhra Pradesh, India
pbhanududi@gmail.com

³ Rao and Naidu College of Engineering, Ongole, Prakasam (Dist), Andhra Pradesh, India
pvnaganjaneyulu@yahoo.co.in

Abstract. Digital watermarking is one of the solutions to protect intellectual properties and copyright by hiding information, such as a random sequence or a logo, into digital media. This paper deals with a new blind robust watermarking technique for embedding secret medical information as a watermark in a host image. As a host image a color medical image is used, and a blue plane goes through the DWT. The diagnosis information of the patient is used as a watermark binary image. The embedding process is done in the LL Sub band of the blue plane of the host image, for DWT a bi-orthogonal wavelet filter is used for corresponding. This can be extended for store more than one watermark information in one medical image by embedding the watermark images in the different sub bands of the host image. The resulting watermark is having good PSNR and robust against different attacks.

Keywords: Watermark, robust, medical image, DWT, PSNR.

1 Introduction

Nowadays, the developments in digital technology have resulted in explosion in the use of digital media products such as image, audio and video. This raises, however, security concerns due to digital multimedia products high vulnerability to the illegal copying, distribution, manipulation, and other attacks. The digital watermarking systems, in literature, have been developed to remedy these security threats. In order to be effective transformation and secure exchange of medical information, it is required to apply digital watermark [1][2], it must be robust and secure, recoverable from a document, provide the original information embedded reliably, and be non-intrusive and also removable by authorized users. In the conventional medical image watermarking techniques the watermarking is fragile in nature. Medical image watermarking is classified into two. 1) Tamper detection and authentication and 2) EPR data hiding. Tamper detection watermarks are able to locate the regions or pixels of the image where tampering was done. Authentication watermarks are used to identify the source of image. EPR data hiding techniques give more importance in

hiding high payload data in the images keeping the imperceptibility very high. Different kinds of watermarking methods were identified for medical images.

This paper deals with the blind and Robust watermarking technique is proposed which are difficult to remove from the object in which they are embedded despite a variety of possible attacks by pirates including compression such as JPEG, scaling and aspect ratio changes, rotation, translation, cropping, row and column removal, addition of noise, filtering, statistical attacks [3]. Describing about the different medical image watermarking techniques are represented in the section 1, A generic wavelet transform based image compression scheme is reviewed in section 2, the description of different topological reordering techniques are represented in section 3, the proposed watermarking technique is discussed in section 4, results are described in the section 5, concluding remarks are given in the section 6.

2 Wavelet Based Image Compression

The widely used unitary transforms for data compression are transformations 2-D images are changed from spatial to frequency transformation such as FFT, DCT, and DWT is a better decomposition capability and having high redundancy than remaining frequency transformations. In DWT, an image can be processed by passing it through an analysis filter bank followed by a decimation operation [4]. This analysis filter bank, which consists of a low pass and a high pass filter at each decomposition stage, is commonly used in image compression. When a signal passes through these filters, it is split into two bands.

The low pass filter, which corresponds to an averaging operation, extracts the coarse information of the image. The high pass filter, which corresponds to a differencing operation, extracts the detail information of the image. The output of the filtering operations is then decimated by two. A 2D-DWT can be accomplished by performing two separate one dimensional transforms. First, the image is filtered along the x dimension using low pass and high pass analysis filters and decimated by two. Low pass filtered coefficients are stored on the left part of the matrix and high pass filtered on the right.

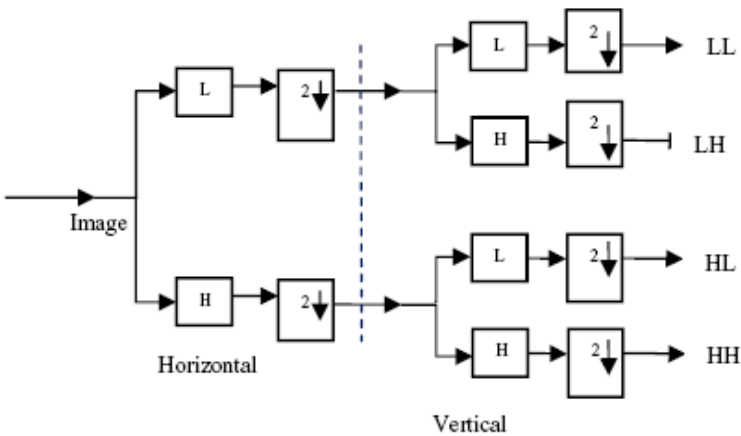


Fig. 1. Pyramidal Decomposition

Because of the decimation, the total size of the transformed image is same as the original image. Then, it is followed by filtering the sub-image along the y-dimension and decimated by two. Finally, the image is split into four bands denoted by LL, HL, LH and HH after one level decomposition and as shown in Figure 1. In the diagram, 'L' indicates low pass filter and 'H' indicates high pass filter. In the second level, the LL band is further decomposed into four bands. The same procedure is continued for further decomposition levels. This process of filtering the image is called pyramidal decomposition of image.

2.1 Bi-orthogonal Wavelet Filter

Where the associated wavelet transforms is invertible but not necessarily orthogonal. Designing bi-orthogonal wavelets allows more degrees of freedoms than orthogonal wavelets [38]. One additional degree of freedom is the possibility to construct symmetric wavelet functions. In the bi-orthogonal case, there are two scaling functions $\phi(t)$, $\overline{\phi}(t)$, which may generate different multiresolution analyses, and accordingly two different wavelet functions instead of just one $\varphi(t), \tilde{\varphi}(t)$

3 Topological Reordering

The rearrangement of coefficients is based purely upon coefficient position, rather than a function of coefficient value, hence it is designated as topological re-ordering. The advantage of re-ordering is better accessibility of the successive coefficients for the estimation of current coefficient value. Due to re-ordering, the size of a particular sub band is altered.

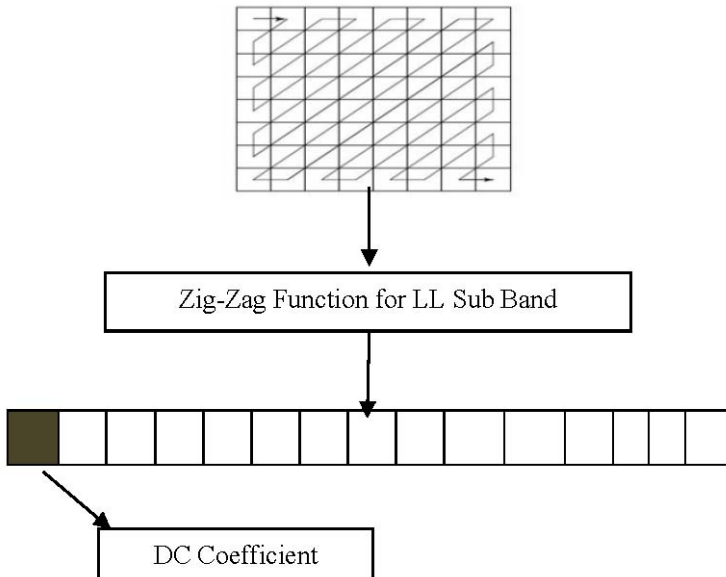


Fig. 2. Performing Zigzag Scanning for the DWT Coefficients of LL Sub band

3.1 Zig-Zag Coding

A zigzag code is described by a highly structured zigzag graph. Due to the structural properties of the graph, it has low-complexity. Zigzag coding is the technique is used to re order the coefficients of the image Sub bands. This process orders the coefficients from low frequency to high frequency [7-11] [13]. The scanning is starts from the top left most Coefficients. This provides better entropy coding and has high access with the successive pixel, the coefficients are get scanned according to the zigzag ordering, to achieve the perceptual invisibility of the mark, without a loss of robustness against signal processing techniques. By Zigzag coding it is possible to access the successive pixels easily. The process of the zigzag is shown in bellow Figure 2. For medical images zigzag scanning is used, where it can possible to scan the medical image without loss in the image pixel. The scanning processes get starts from top to bottom and left to right without neglecting any coefficient. In this work zigzag coding is applied to the LL Sub band of the Blue plane of the Image.

4 Proposed Watermarking Technique

With the proposed watermarking techniques the secrete logo image is embedded in the host image and extracted from the image so the proposed techniques is having both embedding and extraction procedures

4.1 Water Mark Embedding Procedure

This section describes the proposed watermarking technique. The watermark is embedded in the quantized and rounded LL sub band. The embedding algorithm consists of five main steps; the block diagram of the watermark embedding technique is shown in the figure 3.

The Embedding process is shown in the Figure 3. Isolating the image into three color planes. The blue plane information of image is passed through a wavelet transformation unit for the resolution decomposition of the blue plane information. Hierarchical pyramidal decomposition architecture with bank of low pass and high pass filters are used for the decomposition of the original image. The transformation is performed using Bi-orthonormal wavelet transform. Wavelet transform is a very useful tool for signal analysis and image processing, especially in multi-resolution representation. The image is go with the 3 level image decomposition and the coefficients of LL sub band information on which watermark image could be embedded. For the transformation a Zigzag scanning operation is carried out. Similarly for the Logo image also perform the zigzag coding.

The coefficients after zigzag coding the image of 2D is converted to 1D image for the watermarking process the lease significant bits of the image are considered and the watermarking is carried out at the lease significant bits by considering a constant of watermarking strength coefficient.

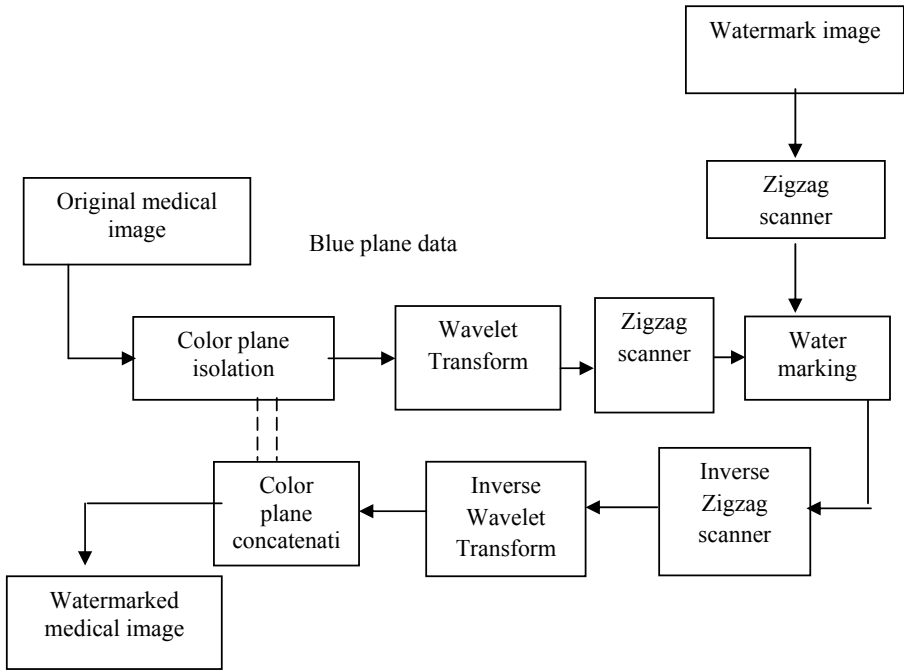


Fig. 3. Proposed watermark embedding system

4.2 Watermark Extraction Procedure

Our watermarking scheme deals with the extraction of the watermark information in the absence of the original image. The aim of the watermark extraction process is to reliably obtain an estimate of the original watermark from a possibly distorted version of the watermarked image. The detection process is inverse procedure of the watermark insertion process. One of the advantages of wavelet-based watermarking is its ability to spread the watermark all over the image. If a part of the image is cropped, it may still contain parts of the watermark. These parts of watermark may be detected by certain mechanism even if the image has been further scaled or rotated.

The Logo that embedded in the medical image is get extracted based on the location where the logo is get placed and the extraction process is shown in the block diagram.

In the extraction algorithm the watermarked image is get separated into the color planes in which the Blue plane goes under the extraction process that is given to the zigzag scanning it convert the 2D Blue plane image into the 1D image that is given to the logo extractor. The location of the watermark that be stored should know to the receiver and the strength of the watermark should be known than only the image is extracted from the Blue plane and that should be in a 1D mode so it need to convert that to 2D Signal the coefficients after the watermark extractor are given to the Inverse Zigzag from this we can extract the logo the block diagram of the Extraction process is shown in the Figure 4.

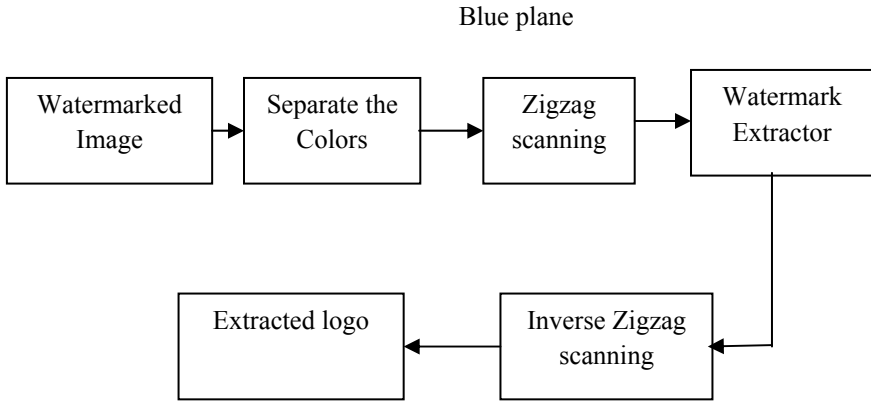


Fig. 4. Proposed watermark extracting system

5 Results

Experiments are performed on five color medical images to verify the proposed method. These five images are represented by 110x110 host image and a logo binary image of 64x64 image is used for the embedding in the host image Here the Image is processed under different attacks as Geometrical, Filtering, and Noise attacks. And also measured the performance, evaluate the quality of the Extracted image. For embedding the logo image in the host image the watermark image is embedded with the strength of 0.5.

5.1 Without Attacks

Experimental results have been obtained with an implementation of the proposed watermarking scheme on ‘medical’ image. Binary logo of size 64x64 is used in the

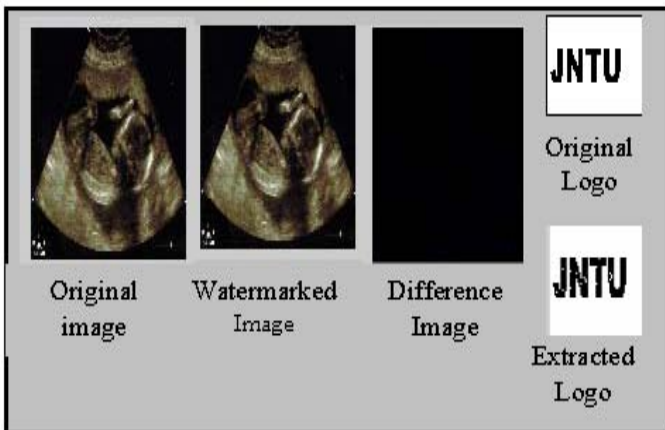









Fig. 5. Watermark image (‘logo’), Watermarked error image, Watermarked image, Extracted Logo.

experiments. 'Medical image' is initially go through the DWT. Reordered watermark is embedded in the re-ordered LL sub band. Watermark bit is embedded at a location of last 4096 bits. Watermark can be embedded in level 2 decomposition also. Original image, original logo, watermarked image and extracted logo are shown in Figure 5.the proposed method has the PSNR value is 39.3dB.

5.2 With Attacks

The proposed method is robust to various noise and filtering attacks. Watermark is extracted even after tampering. The other attacks, where the watermark is successfully retrieved are adding different noises, rotation, sharpening, average filtering and rectangular cropping.

Table 1. Performance of the proposed watermark by different signal processing attacks

S.No	Attack	Quantity	Retrieved watermark	PSNR	NCC	SIM
1	Salt and pepper noise	0.08		30.48	0.975	0.893
2	Sharpening	5x5		31.18	0.832	0.841
3	Gaussian noise	0.05		28.23	0.956	0.832
4	Rotation by angle:	+5°		35.12	0.953	0.849
5	Rectangular Cropping:	50%		20.32	0.832	0.762
6	Average Filtering	9x9		29.14	0.770	.8231
7	Speckle noise	0.08		28.88	0.962	0.892

6 Conclusion

The Proposed watermarking technique is highly robust against different Signal Processing attacks. The algorithm based on the discrete wavelet transformation, with by orthogonal filter, watermark is embedded using Blind watermarking technique.

Hidden watermark image can retrieve from the original image. Thus for watermarking Blue plane of the color image is used, due to blue plane of the image is having less human eye sensitive characteristics. The proposed watermarking is processed under different signal processing attacks; quality of the extracted image is analyzed by PSNR, NCC, and SIM. Topological re-ordering plays major role for protecting the secure data in LL (Low Level) Sub band of the Blue plane image. We can extend this concept by embedding the secrete information in the host image for remaining sub bands of the original image. For the proposed watermarking technique PSNR Value without attacks is 39.3 dB.

References

- [1] Planitz, B., Maeder, A.: e-Health Research Centre. ICT CSIRO Brisbane, QLD 4000 Medical Image Watermarking: A Study on Image Degradation
- [2] Zain, J.: Security in Telemedicine: Watermarking medical images
- [3] Mohidul Islam, S.M., Debnath, R., Alamgir Hossain, S.K.: DWT Based Digital Watermarking Technique and its Robustness on Image Rotation, Scaling, JPEG compression, Cropping and Multiple Watermarking. In: International Conference on Information and Communication Technology (ICICT 2007), Dhaka, Bangladesh, March 7-9 (2007)
- [4] Averbuch, A.Z., Meyer, F., Stromberg, J.O.: Fast adaptive wavelet packet image compression. *IEEE Transactions on Image Processing* 9(8), 792–800 (2000)
- [5] Zhang, L., Gao, W., Wang, Q., Zhao, D.: Macro block-Level Adaptive Scan Scheme for Discrete Cosine Transform Coefficients. In: 2007 IEEE Image Processing (2007)
- [6] Zhou, H., Qi, C., Gao, X.: Low Luminance Smooth Blocks Based Watermarking Scheme in DCT Domain (0-7803-9584-0/06/\$20.00O2006 IEEE)
- [7] Zaidee, A.A., Member IEEE, Fazdliana, S., Member IEEE, Adznan, B.J.: Content Access Control for JPEG Images. In: Proceedings of The Sixth IEEE International Conference on Computer and Information Technology, CIT 2006 (2006)
- [8] Hanbne Trichili, M., Basselsolaiman, o.S.B.: New hybrid scheme of image watermarking Robust vs. jpeg and jpeg.2000 compression. *IEEE, Los Alamitos* (2000)
- [9] Voyatzis, G., Nikolaidis, N., Pitas, I.: Digital watermarking: An overview. In: *EUSIPCO*, vol. 1, pp. 9–12 (1998)
- [10] Veeraswamy, K., SrinivasKumar, S.: An Improved Wavelet based Image Compression Scheme and Oblivious Watermarking. *IJCSNS International Journal of Computer Science and Network Security* 8(4) (April 2008)

Global Chaos Synchronization of Lü and Pan Systems by Adaptive Nonlinear Control

Sundarapandian Vaidyanathan¹ and Karthikeyan Rajagopal²

¹ R & D Centre, Vel Tech Dr. RR & Dr. SR Technical University
Avadi-Alamathi Road, Avadi, Chennai-600 062, India
sundarvtu@gmail.com

<http://www.vel-tech.org/>

² School of Electronics and Electrical Engineering, Singhanian University
Dist. Jhunjhunu, Rajasthan-333 515, India
rkarthikeyan@gmail.com

<http://www.singhanianuniversity.co.in/>

Abstract. This paper investigates the adaptive synchronization of identical Pan systems (2010) and synchronization of non-identical Lü system (2002) and Pan system (2010). In adaptive synchronization of identical chaotic systems, the parameters of the master and slave systems are unknown and we devise feedback control laws using estimates of the system parameters. In adaptive synchronization of different chaotic systems, the parameters of the master system are known, but the parameters of the slave system are unknown and we devise feedback control laws using the estimates of the parameters of the slave system. Our adaptive synchronization results derived in this paper are established using Lyapunov stability theory. Since the Lyapunov exponents are not required for these calculations, the adaptive control method is effective and convenient to synchronize identical and different Lü and Pan systems. Numerical simulations are given to illustrate the effectiveness of the proposed adaptive synchronization schemes for the global chaos synchronization of chaotic systems addressed in this paper.

Keywords: Adaptive control, chaos synchronization, nonlinear control, Lü system, Pan system.

1 Introduction

Chaos theory has been developed and extensively studied over the past two decades. A chaotic system is a nonlinear dynamical system that is highly sensitive to initial conditions. This sensitivity is popularly referred to as the *butterfly effect* [1]. There has been active research done in the past three decades on the control and synchronization of nonlinear chaotic systems.

Chaos synchronization problem was first described by Fujisaka and Yemada [2] in 1983. This problem did not receive great attention until Pecora and Carroll ([3]-[4]) published their results on chaos synchronization in early 1990s. From then on, chaos synchronization has been extensively and intensively studied in the last three decades ([3]-[17]). Chaos theory has been explored in a variety of fields including physical [5], chemical [6], ecological [7] systems, secure communications ([8]-[9]) etc.

In most of the chaos synchronization approaches, the master-slave or drive-response formalism is used. If a particular chaotic system is called the *master* or *drive* system and another chaotic system is called the *slave* or *response* system, then the idea of the synchronization is to use the output of the master system to control the slave system so that the output of the slave system tracks the output of the master system asymptotically.

Since the seminal work by Pecora and Carroll ([3]-[4]), a variety of impressive approaches have been proposed for the synchronization for the chaotic systems such as PC method ([3]-[4]), the sampled-data feedback synchronization method [10], OGY method [11], time-delay feedback approach [12], backstepping design method [13], adaptive design method [14], sliding mode control method [15], active control method ([16]-[17]), etc.

In this paper, we apply adaptive nonlinear control method to derive new results for the global chaos synchronization for identical Pan systems ([18], 2010) and non-identical Lü system ([19], 2002) and Pan system.

In adaptive synchronization of identical chaotic systems, the parameters of the master and slave systems are unknown and we devise feedback control laws using estimates of the system parameters. In adaptive synchronization of different chaotic systems, the parameters of the master system are known, but the parameters of the slave system are unknown and we devise feedback control laws using the estimates of the parameters of the slave system.

This paper has been organized as follows. In Section 2, we derive results for the adaptive synchronization of identical Pan systems. In Section 3, we derive results for the adaptive synchronization of non-identical Lü and Pan systems. In Section 4, we summarize the main results obtained in this paper.

2 Adaptive Synchronization of Identical Pan Systems

2.1 Theoretical Results

In this section, we discuss the adaptive synchronization of identical Pan systems (Pan, Xu and Zhou, [18], 2010), where the parameters of the master and slave systems are unknown.

As the master system, we consider the Pan dynamics described by

$$\begin{aligned}\dot{x}_1 &= \alpha(x_2 - x_1) \\ \dot{x}_2 &= \gamma x_1 - x_1 x_3 \\ \dot{x}_3 &= x_1 x_2 - \beta x_3\end{aligned}\tag{1}$$

where x_1, x_2, x_3 are the state variables and α, β, γ are unknown parameters of the system.

The Pan system (1) is chaotic when the parameter values are

$$\alpha = 10, \quad \beta = 8/3 \quad \text{and} \quad \gamma = 16$$

The chaotic state portrait of the Pan system (1) is shown in Figure 1

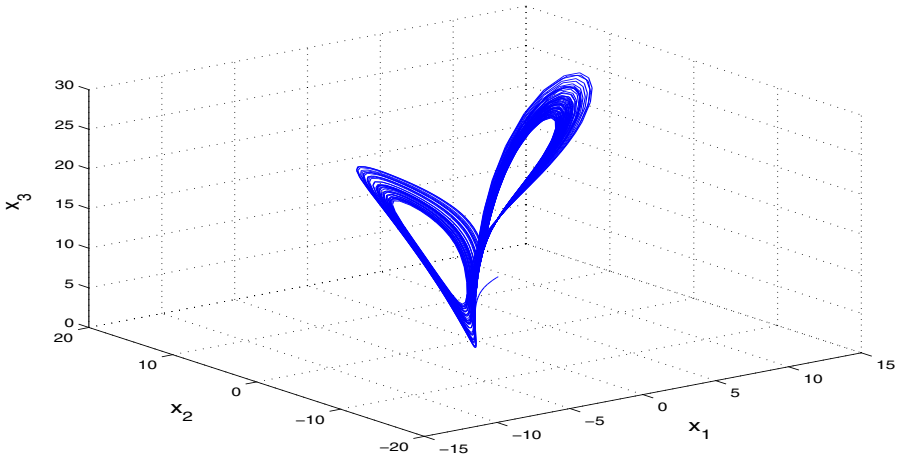


Fig. 1. State Orbits of the Pan System

As the slave system, we consider again the Pan dynamics described by

$$\begin{aligned}\dot{y}_1 &= \alpha(y_2 - y_1) + u_1 \\ \dot{y}_2 &= \gamma y_1 - y_1 y_3 + u_2 \\ \dot{y}_3 &= y_1 y_2 - \beta y_3 + u_3\end{aligned}\quad (2)$$

where y_1, y_2, y_3 are the state variables, α, β, γ are unknown parameters of the system and u_1, u_2, u_3 are the nonlinear controls to be designed.

The synchronization error is defined by

$$e_i = y_i - x_i, \quad (i = 1, 2, 3)$$

Then the error dynamics is obtained as

$$\begin{aligned}\dot{e}_1 &= \alpha(e_2 - e_1) + u_1 \\ \dot{e}_2 &= \gamma e_1 - y_1 y_3 + x_1 x_3 + u_2 \\ \dot{e}_3 &= y_1 y_2 - x_1 x_2 - \beta e_3\end{aligned}\quad (3)$$

Let us now define the adaptive control functions $u_1(t), u_2(t), u_3(t)$ as

$$\begin{aligned}u_1(t) &= -\hat{\alpha}(e_2 - e_1) - k_1 e_1 \\ u_2(t) &= -\hat{\gamma} e_1 + y_1 y_3 - x_1 x_3 - k_2 e_2 \\ u_3(t) &= -y_1 y_2 + x_1 x_2 + \hat{\beta} e_3 - k_3 e_3\end{aligned}\quad (4)$$

where $\hat{\alpha}, \hat{\beta}$ and $\hat{\gamma}$ are estimates of α, β and γ respectively, and $k_i, (i = 1, 2, 3)$ are positive constants.

Let us now define the parameter errors as

$$e_\alpha = \alpha - \hat{\alpha}, \quad e_\beta = \beta - \hat{\beta} \quad \text{and} \quad e_\gamma = \gamma - \hat{\gamma}\quad (5)$$

Substituting (4) and (5) into (3), the error dynamics can be obtained as

$$\begin{aligned} \dot{e}_1 &= e_\alpha (e_2 - e_1) - k_1 e_1 \\ \dot{e}_2 &= -e_\gamma e_1 - k_2 e_2 \\ \dot{e}_3 &= -e_\beta e_3 - k_3 e_3 \end{aligned} \tag{6}$$

For the derivation of the update law for adjusting the estimates of the parameters, the Lyapunov approach is used. Consider the quadratic Lyapunov function

$$V(e_1, e_2, e_3, e_\alpha, e_\beta, e_\gamma) = \frac{1}{2} (e_1^2 + e_2^2 + e_3^2 + e_\alpha^2 + e_\beta^2 + e_\gamma^2) \tag{7}$$

Note also that

$$e_{\dot{\alpha}} = -\dot{\hat{\alpha}}, \quad e_{\dot{\beta}} = -\dot{\hat{\beta}} \quad \text{and} \quad e_{\dot{\gamma}} = -\dot{\hat{\gamma}} \tag{8}$$

Differentiating V using (6) and (8), we get

$$\dot{V} = -k_1 e_1^2 - k_2 e_2^2 - k_3 e_3^2 + e_\alpha [e_1(e_2 - e_1) - \dot{\hat{\alpha}}] + e_\beta [-e_3^2 - \dot{\hat{\beta}}] + e_\gamma [-e_1 e_2 - \dot{\hat{\gamma}}] \tag{9}$$

In view of Eq. (9), the estimated parameters are updated by the following law:

$$\begin{aligned} \dot{\hat{\alpha}} &= e_1(e_2 - e_1) + k_4 e_\alpha \\ \dot{\hat{\beta}} &= e_3^2 + k_5 e_\beta \\ \dot{\hat{\gamma}} &= e_1 e_2 + k_6 e_\gamma \end{aligned} \tag{10}$$

where k_4, k_5 and k_6 are positive constants.

Substituting (10) into (9), we get

$$\dot{V} = -k_1 e_1^2 - k_2 e_2^2 - k_3 e_3^2 - k_4 e_\alpha^2 - k_5 e_\beta^2 - k_6 e_\gamma^2 \tag{11}$$

which is a negative definite function.

Thus, by Lyapunov stability theory [20], it is immediate that the synchronization error and parameter error decay to zero exponentially with time.

Hence, we have proved the following result.

Theorem 1. *The identical Pan systems (1) and (2) with unknown parameters are globally and exponentially synchronized by the adaptive control law (4), where the update law for the parameters is given by (10) and $k_i, (i = 1, \dots, 6)$ are positive constants. ■*

2.2 Numerical Results

For the numerical simulations, the fourth-order Runge-Kutta method is used to solve the systems using MATLAB.

We consider the adaptive synchronization of the Pan systems with $\alpha = 10, \beta = 8/3$ and $\gamma = 16$. Suppose that the initial estimated parameters are $\hat{\alpha}(0) = 5, \hat{\beta}(0) = 4$ and $\hat{\gamma}(0) = 2$. We apply the adaptive control law (4) and the update law (10) with $k_i = 2, (i = 1, \dots, 6)$.

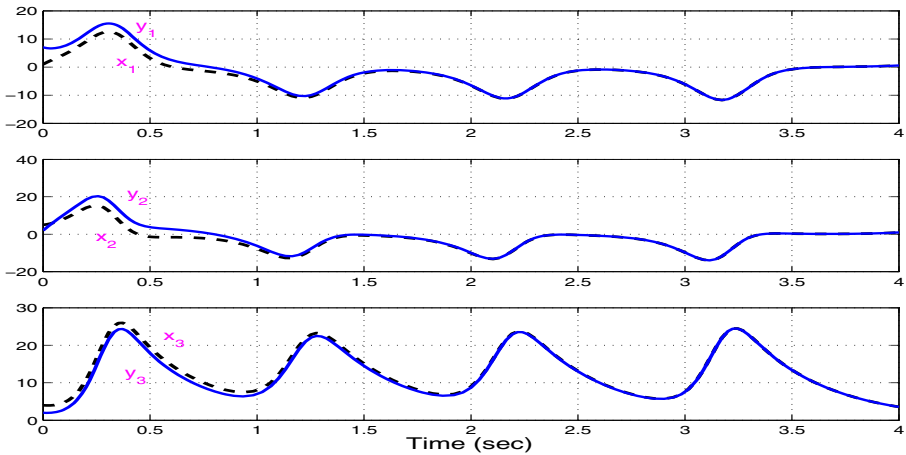


Fig. 2. Adaptive Synchronization of the Identical Pan Systems

The initial conditions of the master system (1) are taken as $x(0) = (1, 5, 4)$ and the initial conditions of the slave system (2) are taken as $y(0) = (7, 2, 2)$.

Figure 2 shows the complete synchronization of the identical Pan systems and Figure 3 shows that the estimated values of parameters $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\gamma}$ converge to the system parameters $\alpha = 10$, $\beta = 8/3$ and $\gamma = 16$ respectively.

3 Adaptive Synchronization of Non-identical Lü and Pan Systems

3.1 Theoretical Results

In this section, we discuss the adaptive synchronization of non-identical Lü and Pan systems. Here, we consider the Lü system ([19], 2002) as the master system, whose parameters are known. We consider the Pan system ([18], 2010) as the slave system, whose parameters are unknown. The problem is to devise adaptive feedback control law and parameter update law which synchronize the master and slave systems.

As the master system, we consider the Lü dynamics described by

$$\begin{aligned} \dot{x}_1 &= a(x_2 - x_1) \\ \dot{x}_2 &= -x_1x_3 + cx_2 \\ \dot{x}_3 &= x_1x_2 - bx_3 \end{aligned} \tag{12}$$

where x_1, x_2, x_3 are the state variables and a, b, c are known, positive, constant parameters of the system.

The chaotic state portrait of the Lü system (12) is shown in Figure 4.

As the slave system, we consider the Pan dynamics described by

$$\begin{aligned} \dot{y}_1 &= \alpha(y_2 - y_1) + u_1 \\ \dot{y}_2 &= \gamma y_1 - y_1y_3 + u_2 \\ \dot{y}_3 &= y_1y_2 - \beta y_3 + u_3 \end{aligned} \tag{13}$$

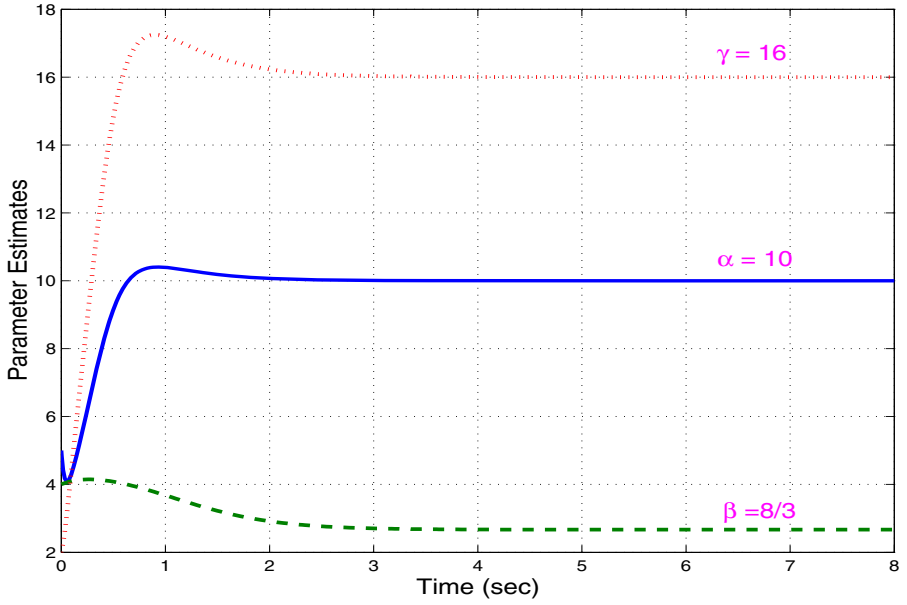


Fig. 3. Parameter Estimates $\hat{\alpha}(t), \hat{\beta}(t), \hat{\gamma}(t)$

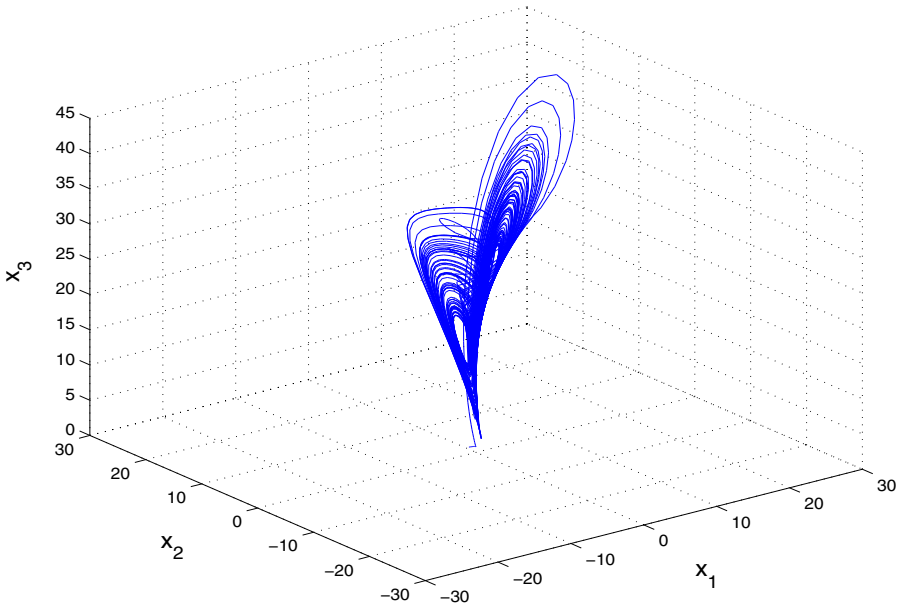


Fig. 4. State Orbits of the Lü System

where y_1, y_2, y_3 are the state variables, α, β, γ are unknown parameters of the system and u_1, u_2, u_3 are the nonlinear controls to be designed.

The synchronization error is defined by

$$e_i = y_i - x_i, \quad (i = 1, 2, 3)$$

Then the error dynamics is obtained as

$$\begin{aligned} \dot{e}_1 &= \alpha(y_2 - y_1) - a(x_2 - x_1) + u_1 \\ \dot{e}_2 &= \gamma y_1 - cx_2 - y_1 y_3 + x_1 x_3 + u_2 \\ \dot{e}_3 &= -\beta y_3 + bx_3 + y_1 y_2 - x_1 x_2 + u_3 \end{aligned} \tag{14}$$

Let us now define the adaptive control functions $u_1(t), u_2(t), u_3(t)$ as

$$\begin{aligned} u_1(t) &= -\hat{\alpha}(y_2 - y_1) + a(x_2 - x_1) - k_1 e_1 \\ u_2(t) &= -\hat{\gamma} y_1 + cx_2 + y_1 y_3 - x_1 x_3 - k_2 e_2 \\ u_3(t) &= \hat{\beta} y_3 - bx_3 - y_1 y_2 + x_1 x_2 - k_3 e_3 \end{aligned} \tag{15}$$

where $\hat{\alpha}, \hat{\beta}$ and $\hat{\gamma}$ are estimates of α, β and γ respectively, and $k_i, (i = 1, 2, 3)$ are positive constants.

Let us now define the parameter errors as

$$e_\alpha = \alpha - \hat{\alpha}, \quad e_\beta = \beta - \hat{\beta} \quad \text{and} \quad e_\gamma = \gamma - \hat{\gamma} \tag{16}$$

Substituting (15) and (16) into (14), the error dynamics can be obtained as

$$\begin{aligned} \dot{e}_1 &= e_\alpha (y_2 - y_1) - k_1 e_1 \\ \dot{e}_2 &= e_\gamma y_1 - k_2 e_2 \\ \dot{e}_3 &= -e_\beta y_3 - k_3 e_3 \end{aligned} \tag{17}$$

For the derivation of the update law for adjusting the estimates of the parameters, the Lyapunov approach is used.

Consider a quadratic function

$$V(e_1, e_2, e_3, e_\alpha, e_\beta, e_\gamma) = \frac{1}{2} (e_1^2 + e_2^2 + e_3^2 + e_\alpha^2 + e_\beta^2 + e_\gamma^2) \tag{18}$$

Note also that

$$\dot{e}_\alpha = -\dot{\hat{\alpha}}, \quad \dot{e}_\beta = -\dot{\hat{\beta}} \quad \text{and} \quad \dot{e}_\gamma = -\dot{\hat{\gamma}} \tag{19}$$

Differentiating V using (17) and (19), we get

$$\dot{V} = -k_1 e_1^2 - k_2 e_2^2 - k_3 e_3^2 + e_\alpha [e_1(y_2 - y_1) - \dot{\hat{\alpha}}] + e_\beta [-e_3 y_3 - \dot{\hat{\beta}}] + e_\gamma [e_2 y_1 - \dot{\hat{\gamma}}] \tag{20}$$

In view of Eq. (20), the estimated parameters are updated by the following law:

$$\begin{aligned} \dot{\hat{\alpha}} &= e_1(y_2 - y_1) + k_4 e_\alpha \\ \dot{\hat{\beta}} &= -e_3 y_3 + k_5 e_\beta \\ \dot{\hat{\gamma}} &= e_2 y_1 + k_6 e_\gamma \end{aligned} \tag{21}$$

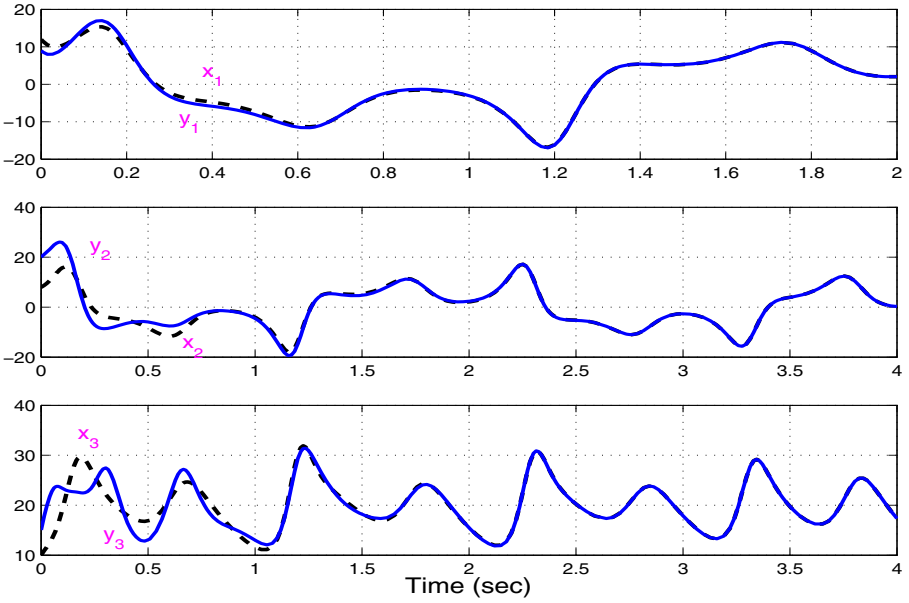


Fig. 5. Adaptive Synchronization of the Lü and Pan Systems

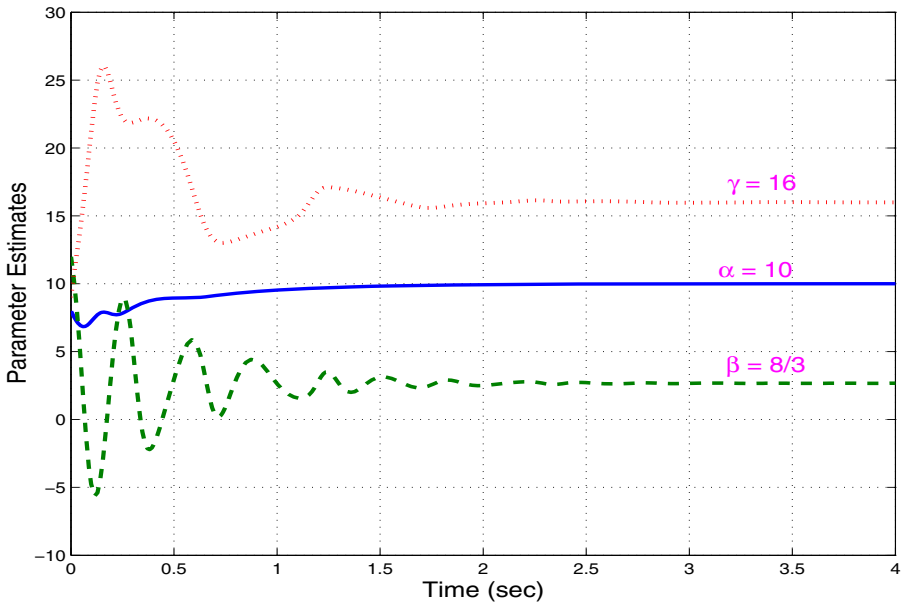


Fig. 6. Parameter Estimates $\hat{\alpha}(t), \hat{\beta}(t), \hat{\gamma}(t)$

where k_4, k_5 and k_6 are positive constants.

Substituting (21) into (20), we get

$$\dot{V} = -k_1 e_1^2 - k_2 e_2^2 - k_3 e_3^2 - k_4 e_\alpha^2 - k_5 e_\beta^2 - k_6 e_\gamma^2 \quad (22)$$

which is a negative definite function.

Thus, by Lyapunov stability theory [20], it is immediate that the synchronization error and parameter error decay to zero exponentially with time.

Hence, we have proved the following result.

Theorem 2. *The non-identical Lü system (12) with known parameters and Pan system (13) with unknown parameters are globally and exponentially synchronized by the adaptive control law (15), where the update law for the parameters is given by (21) and $k_i, (i = 1, \dots, 6)$ are positive constants. ■*

3.2 Numerical Results

For the numerical simulations, the fourth-order Runge-Kutta method is used to solve the systems using MATLAB.

We consider the adaptive synchronization of Lü and Pan systems with

$$a = 36, \quad b = 3, \quad c = 20, \quad \alpha = 10, \quad \beta = 8/3 \quad \text{and} \quad \gamma = 16$$

Suppose that the initial estimated parameters of the slave system are

$$\hat{\alpha}(0) = 8, \quad \hat{\beta}(0) = 12 \quad \text{and} \quad \hat{\gamma}(0) = 9$$

We apply the adaptive control law (15) and the update law (21) with

$$k_1 = 2, \quad k_2 = 2, \quad k_3 = 2, \quad k_4 = 2, \quad k_5 = 2, \quad k_6 = 2$$

The initial conditions of the master system (12) are taken as $x(0) = (12, 8, 10)$ and the and the initial conditions of the slave system (13) are taken as $y(0) = (9, 20, 15)$.

Figure 5 shows the complete synchronization of the non-identical Lü and Pan systems and Figure 6 shows that the estimated values of parameters $\hat{\alpha}, \hat{\beta}$ and $\hat{\gamma}$ converge to the slave system parameters $\alpha = 10, \beta = 8/3$ and $\gamma = 16$ respectively.

4 Conclusions

In this paper, we have used adaptive control method to achieve global chaos synchronization of identical Pan systems (2010) and non-identical Lü system (2002) and Pan system. The adaptive synchronization results derived in this paper are established using Lyapunov stability theory. Since the Lyapunov exponents are not needed for these calculations, the proposed adaptive control method is very effective and convenient for synchronizing chaotic systems. Numerical simulations are also shown to demonstrate the effectiveness of the adaptive synchronization schemes derived in this paper.

References

1. Alligood, K.T., Sauer, T., Yorke, J.A.: *Chaos: An Introduction to Dynamical Systems*. Springer, New York (1997)
2. Fujikasa, H., Yamada, T.: Stability theory of synchronized motion in coupled-oscillator systems. *Progr. Theoret. Phys.* 69, 32–47 (1983)
3. Pecora, L.M., Carroll, T.L.: Synchronization in chaotic systems. *Phys. Rev. Lett.* 64, 821–824 (1990)
4. Pecora, L.M., Carroll, T.L.: Synchronizing chaotic circuits. *IEEE Trans. Circ. Sys.* 38, 453–456 (1991)
5. Lakshmanan, M., Murali, K.: *Chaos in Nonlinear Oscillators: Controlling and Synchronization*. World Scientific, Singapore (1996)
6. Han, S.K., Kerrer, C., Kuramoto, Y.: Dephasing and bursting in coupled neural oscillators. *Phys. Rev. Lett.* 75, 3190–3193 (1995)
7. Blasius, B., Huppert, A., Stone, L.: Complex dynamics and phase synchronization in spatially extended ecological system. *Nature* 399, 354–359 (1999)
8. Kwok, H.S., Wallace, K., Tang, S., Man, K.F.: Online secure communication system using chaotic map. *Internat. J. Bifurcat. Chaos* 14, 285–292 (2004)
9. Kocarev, L., Parlitz, U.: General approach for chaos synchronization with applications to communications. *Phys. Rev. Lett.* 74, 5028–5030 (1995)
10. Yang, T., Chua, L.O.: Control of chaos using sampled-data feedback control. *Internat. J. Bifurcat. Chaos* 9, 215–219 (1999)
11. Ott, E., Grebogi, C., Yorke, J.A.: Controlling chaos. *Phys. Rev. Lett.* 64, 1196–1199 (1990)
12. Park, J.H., Kwon, O.M.: A novel criterion for delayed feedback control of time-delay chaotic systems. *Chaos, Solit. Fract.* 17, 709–716 (2003)
13. Wu, X., Lü, J.: Parameter identification and backstepping control of uncertain Lü system. *Chaos, Solit. Fract.* 18, 721–729 (2003)
14. Park, J.H.: Adaptive control for modified projective synchronization of a four-dimensional chaotic system with uncertain parameters. *J. Comput. Applied Math.* 213, 288–293 (2008)
15. Yau, H.T.: Design of adaptive sliding mode controller for chaos synchronization with uncertainties. *Chaos, Solit. Fract.* 22, 341–347 (2004)
16. Huang, L., Feng, R., Wang, M.: Synchronization of chaotic systems via nonlinear control. *Phys. Lett. A* 320, 271–275 (2004)
17. Chen, H.K.: Global chaos synchronization of new chaotic systems via nonlinear control. *Chaos, Solit. Fract.* 23, 1245–1251 (2005)
18. Pan, L., Xu, D., Zhou, W.: Controlling a novel chaotic attractor using linear feedback. *J. Inform. Comput. Sci.* 5, 117–124 (2010)
19. Lü, J., Chen, G.: A new chaotic attractor coined. *Internat. J. Bifur. Chaos* 12, 659–661 (2002)
20. Hahn, W.: *The Stability of Motion*. Springer, New York (1967)

Reducing False Positives in Minutia Detection by Using the Proposed Fingerprint Alignment Technique

P. Jaganathan¹ and M. Rajinikannan²

¹ Department of Computer Applications, P.S.N.A College of Engineering and Technology, Dindigul, Tamilnadu-624622, India
jaganathodc@gmail.com

² Department of Computer Applications, P.S.N.A College of Engineering and Technology, Dindigul, Tamilnadu-624622, India
rajinikannan@gmail.com

Abstract. The important step in automatic fingerprint recognition system is to automatically and reliably extract minutia from the input fingerprint images. In such recognition systems, the orientation of the fingerprint image has influence on fingerprint image enhancement phase, minutia detection phase and minutia matching phase of the system. The fingerprint image rotation, translation, and registration are the commonly used techniques, to minimize the error in all these stages of fingerprint recognition. In this work, we approached two methods by which the minutia is detected. In the first method, the minutias are detected without aligning the image. In the second method, the input image is aligned using the proposed k-means clustering based fingerprint image rotation algorithm and then the minutias are detected. This proposed rotation algorithm could be applied as a pre-processing step before minutia detection. In both the methods the images are enhanced using the proposed Gabor filter. Finally the results clearly show that the aligned images give more accurate true minutias than the unaligned images. Hence, the result will be better detection of minutia as well as better matching with improved performance.

Keywords: Fingerprint Image Enhancement, Ridge Endings, Ridge Bifurcation, Gabor Filter, k-means clustering, Fingerprint Image Rotation, Alignment, Minutia Detection.

1 Introduction

Fingerprint recognition is part of the larger field of Biometrics and has found an extensive application field. Other techniques of biometrics include face recognition, voice recognition, hand geometry, retinal scan, and ear surface and so on. In this context, the minutia extraction plays an important role in the accurate recognition of the fingerprint recognizing system.

1.1 The Fingerprint Image and Minutia

A fingerprint is the feature pattern of one finger. Fingerprints are believed to be unique across individuals, and across fingers of the same individual [7]. Fingerprint-

based personal identification has been used for a very long time [5]. It is believed that each person has his own fingerprints with the permanent uniqueness. Hence, fingerprints are being used for identification and forensic investigation for a long time. Minutia points occur either at a ridge ending or a ridge bifurcation. A ridge ending is defined as the point where the ridge ends abruptly and the ridge bifurcation is the point where the ridge splits into two or more branches. A fingerprint expert is often able to correctly identify the minutia by using various visual clues such as local ridge orientation, ridge continuity, ridge tendency, etc., as long as the ridge and furrow structures are not corrupted completely [1].

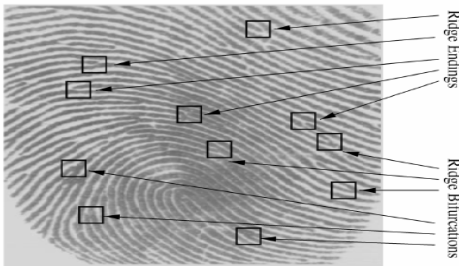


Fig. 1. The Ridge Endings and Bifurcation

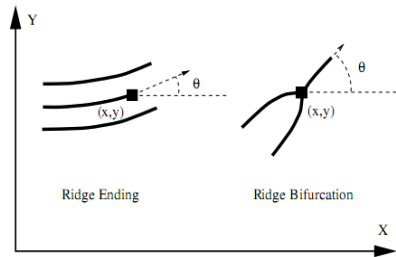


Fig. 2. The Characteristic of a Minutia.

The above diagram shows the clear view of a minutia and its characteristic attributes.

1.2 The Importance of Fingerprint Image Alignment Algorithms

A typical fingerprint recognition system, have several types of algorithms in each stage. We can broadly reduce these stages of algorithms as Fingerprint Image Enhancement (Ridge and Valley Enhancement) Algorithms, Minutia Detection Algorithms and Minutia Matching Algorithms.

Fingerprint Image alignment algorithm is the most important step, since the results of Minutia extraction Algorithms and the Minutia matching Algorithms are very much depend on the clarity and orientation of the fingerprint image. Point pattern matching is generally intractable due to the lack of knowledge about the correspondence between two point's sets. To address this problem, Jain et al. proposed alignment-based minutia matching [3], [2]. Two sets of minutia are first aligned using corresponding ridges to find a reference minutia pair, one from the input fingerprint and another from the template fingerprint, and then all the other minutia of both images are converted with respect to the reference minutia. Jiang et al. also proposed an improved method [4], using minutia together with its neighbors to find the best alignment.

In general, minutia matcher chooses any two minutias as a reference minutia pair and then matches their associated ridges first. If the ridges match well [9], two fingerprint images are aligned and matching is conducted for all remaining minutia. In any of this matching policy, the fingerprint image alignment will play significant role

and will have influence on accuracy. In the following sections, the proposed fingerprint alignment and enhancement algorithms are discussed.

The rest of the paper is organized as follows: Section 2 deals with proposed image alignment and fingerprint image enhancement algorithms. Section 3 deals with Implementation and result analysis. Sections 4 and 5 deals with conclusion and references.

2 The Proposed Image Alignment and Enhancement Algorithm

2.1 Pre-processing the Input Fingerprint Image

Initially, the input fingerprint image is pre-processed by changing the intensity value to 0 for the pixel, whose intensity value is less than 32. Now, only the ridges and valleys of the fingerprint image are considered. Except ridges and valleys the other parts of the fingerprint image pixels intensity values are changed to 0. Finally the binary image will be constructed. The co-ordinates of (x,y) the one valued pixels are considered as two attributes and passed as Input to the K-Means algorithm.

2.2 K-Means Algorithm

K-Means algorithm is very popular for data clustering. Generally, K-Means algorithm is used in several iterations to cluster the data since the result is very much depend on the initial guess of the cluster centres. The Algorithm goes like this

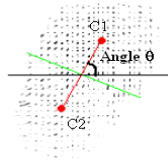
1. Start iteration
2. Select k Center in the problem space (it can be random).
3. Partition the data into k clusters by grouping points that are closest to that k centers.
4. Use the mean of these k clusters to find new centers.
5. Repeat steps 3 and 4 until centers do not change, for N iterations.
6. Among the N results, find the result with minimum distance.
7. Display the results corresponding to that minimum distance.
8. Stop iterations.

2.3 K-Means Algorithm for Fingerprint Image Rotation

Let us consider the pixels of the two-dimensional finger print image as the plotted three data points of x , y , Gray level.



1. The fingerprint image pixels are assumed as data points in 2d space.
2. The data points were clustered in to two groups using k-means clustering algorithm. The green line showing the cluster boundary.



3. The points C1 and C2 are the centers of the two clusters.
4. A line connecting C1 and C2 will be almost equal to the inclination of the fingerprint image.
5. The inclination angle θ can be measured from the base of the image.
 - $\theta = \text{atan} ((x1-x2) / (y1-y2))$ (in radians)
 - $\theta = \theta * (180/\pi)$ (in degree)
 - if $\theta < 0$
 - $\theta = 90 + \theta$
 - else
 - $\theta = - (90 - \theta)$
 - end
6. Now rotating the image by angle θ .
7. The direction of rotation can be decided with respect to the location of the point C1 in the top two quadrants of the four Quadrants.
8. This will finally give the well-aligned image.

2.4 The Gabor Filter Based Fingerprint Image Enhancement

The fingerprint enhancement techniques proposed by Jain [8], is based on the convolution of the image with Gabor filters, which has the local ridge orientation and ridge frequency. The algorithm includes normalization, ridge orientation estimation, ridge frequency estimation and filtering. Gabor filters are band pass filters that have both frequency- selective and orientation-selective properties [10], thus the ridge are enhanced. Gabor filters [11] have both frequency selective and orientation selective properties and have optimal joint resolution in both spatial and frequency domains. Therefore, it is appropriate to use Gabor filters as band-pass filters to remove the noise and preserve true ridge/valley structures The Gabor filter only filters along the specified orientation this decreases anything oriented differently. Hence, the filter increases the contrast and reduces the noise.

$$G(x, y; \theta, f) = \exp \left\{ -\frac{1}{2} \left[\frac{x_{\theta}^2}{\sigma_x^2} + \frac{y_{\theta}^2}{\sigma_y^2} \right] \right\} \cos(2\pi f x_{\theta}), \tag{1}$$

$$x_{\theta} = x \cos \theta + y \sin \theta,$$

$$y_{\theta} = -x \sin \theta + y \cos \theta,$$

In equation (1), θ is the local orientation, f is the frequency, σ_x and σ_y are the standard deviation of the Gaussian Envelope. The convolution will result in the noise being reduced and the ridges being enhanced, since the Gabor filter only filters along

the specified orientation this decreases anything oriented differently. Hence, the filter increases the contrast and reduces the noise. The next step after enhancement of the image is the extraction of minutia. The enhanced image is binarized first. The skeleton of the image is then formed. The binary image is thinned as a result of which a ridge is only one pixel wide. The minutia points are thus those which have a pixel value of one (ridge ending) as their neighbour or more than two ones (ridge bifurcations) in their neighbourhood.

3 Implementation Results and Analysis

The figure 3 shows the evaluation strategy used in this work.

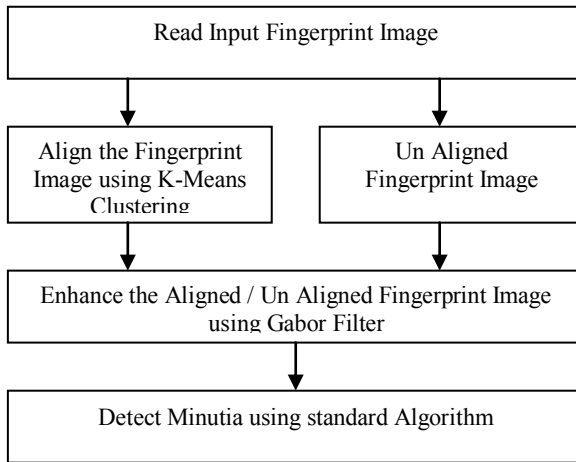


Fig. 3. The Evaluation Model

3.1 Fingerprint Database Used For Evaluation

A fingerprint database from the FVC2000 [6] (Fingerprint Verification Competition 2000) is used to experiment performance of the proposed alignment algorithm. FVC2000 was the First International Competition for Fingerprint Verification Algorithms. This initiative is organized by D. Maio, D. Maltoni, R. Cappelli from Biometric Systems Lab (University of Bologna), J. L. Wayman from the U.S. National Biometric Test Center (San Jose State University) and A. K. Jain from the Pattern Recognition and Image Processing Laboratory of Michigan State University. Sets of few selected images from the above database were used to evaluate the performance of the algorithms.

3.2 Sample Set of Results

The following outputs show some of the inputs as well as the corresponding outputs. In the second column rotated images and the angle of rotation estimated by the k-means based algorithm also given below.





	Input Image	Vertically Aligned Image
102_3.TIF		 k-means : $\theta = 26.5651$
105_2.TIF		 k-means : $\theta = -14.0362$

Fig. 4. The Results of Rotation

Table 1. The total number of ENDS and BRANCHES of Unaligned and Aligned Images using GABOR Enhancement Technique

Image	Un Aligned Fingerprint Image		Aligned Fingerprint Image	
	ENDS	BRANCHES	ENDS	BRANCHES
102_3	57	54	49	41
103_1	42	15	66	24
103_2	54	31	81	29
105_2	99	14	104	22
106_3	97	54	87	43
102_1	35	28	128	23
102_4	104	23	128	23
102_5	87	33	130	36
102_6	89	29	129	40
103_3	27	31	30	30
108_1	122	27	146	29
108_3	138	27	219	54
108_4	107	27	140	31
105_7	19	19	52	29
105_8	47	14	87	49
105_3	60	36	76	58

The following table shows the results of various images total ridge endings (ENDS) and ridge bifurcation (BRANCHES) before Alignment and after Alignment. The Gabor filter is used for Image enhancement and standard Minutia detection algorithm is used to count the ridge endings and bifurcation.

In the following table, before alignment, there are many falsely detected branches around the edges of the image. But after alignment the false detections were considerably reduced.

Table 2. After alignment the false detection reduced

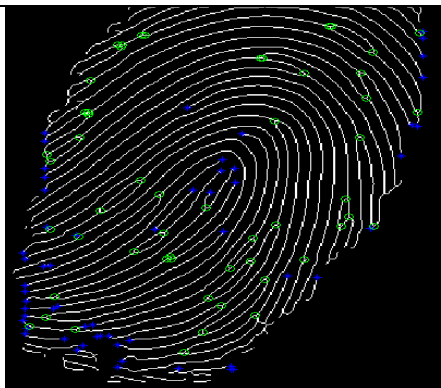
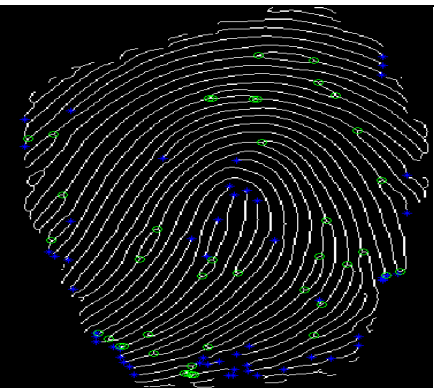
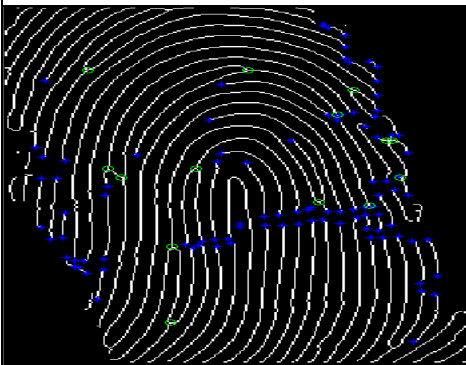
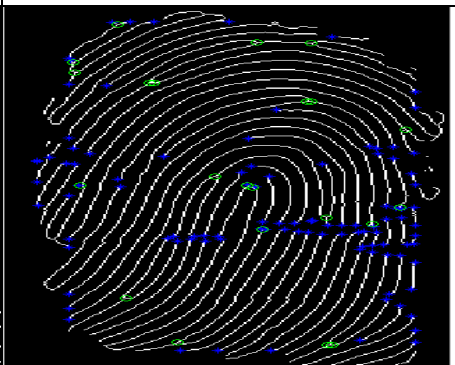
Before Alignment	After Alignment
 <p data-bbox="296 825 415 853" style="text-align: center;">102_3.TIF</p>	 <p data-bbox="759 825 879 853" style="text-align: center;">102_3.TIF</p>
<p data-bbox="158 860 560 889" style="text-align: center;">The Total Number of Branches: 54</p> <p data-bbox="181 889 537 917" style="text-align: center;">The Total Number of Ends: 57</p>	<p data-bbox="622 860 1024 889" style="text-align: center;">The Total Number of Branches: 41</p> <p data-bbox="645 889 1001 917" style="text-align: center;">The Total Number of Ends: 49</p>

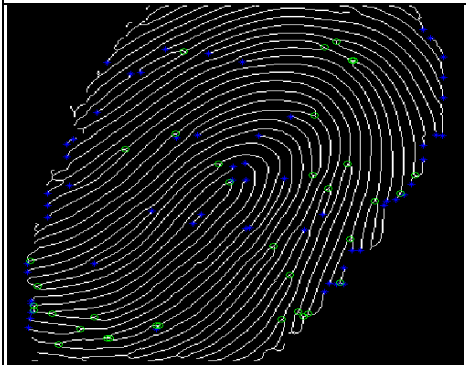
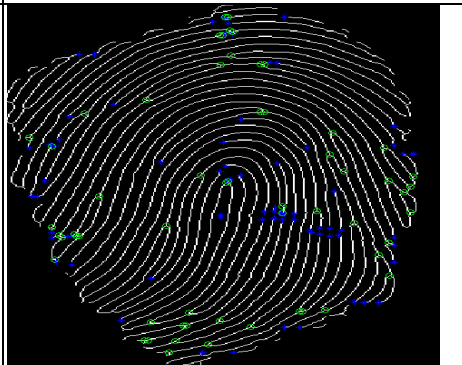
Table 3. After alignment more true branches are detected

Before Alignment	After Alignment
 <p data-bbox="296 1508 415 1536" style="text-align: center;">105_2.TIF</p>	 <p data-bbox="759 1508 879 1536" style="text-align: center;">105_2.TIF</p>
<p data-bbox="158 1543 560 1571" style="text-align: center;">The Total Number of Branches: 14</p> <p data-bbox="181 1571 537 1599" style="text-align: center;">The Total Number of Ends: 99</p>	<p data-bbox="594 1543 996 1571" style="text-align: center;">The Total Number of Branches: 22</p> <p data-bbox="617 1571 973 1599" style="text-align: center;">The Total Number of Ends: 104</p>

In the following table, before alignment, there are only minimum truly detected branches in the image. But more number of branches is detected after alignment.

In the following table, before alignment, there is only minimum number of truly detected branches in the image. If you see the bottom side of the image there are too many branches that were not detected correctly. But more significant branches are detected after alignment.

Table 4. After alignment more significant branches are detected

Before Alignment	After Alignment
	
105_3.TIF	105_3.TIF
<p style="text-align: center;">The Total Number of Branches: 36 The Total Number of Ends: 60</p>	<p style="text-align: center;">The Total Number of Branches: 58 The Total Number of Ends: 76</p>

4 Conclusion and Future Work

We have successfully implemented and evaluated the proposed k-means based fingerprint image alignment and Gabor Filter based Image Enhancement algorithm with Matlab 6.5. The arrived results clearly show that the better minutia is detected only after fingerprint image alignment and also it reduces the false positives. Hence, if we use the aligned fingerprint image for the detection of minutia and minutia matching, then we may expect better accuracy in recognition. Future works may evaluate the difference in recognition with and without the proposed fingerprint image alignment phase.

References

1. Hong, L., Wan, Y., Jain, A.K.: Fingerprint image enhancement: Algorithm and performance evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 777–789 (1998)

2. Jain, A.K., Hong, L., Pankanti, S., Bolle, R.: An Identity authentication system using fingerprints. *Proc. IEEE* 85(9), 1365–1388 (1997)
3. Jain, A.K., Hong, L., Pankanti, S., Bolle, R.: On-Line fingerprint verification. *IEEE Trans. Pattern Anal. Mach. Intell.* 19(4), 302–314 (1997)
4. Jiang, X., Yau, W.: Fingerprint minutia matching based on the local and global structures. In: *Proc. 15th Int. Conf. Pattern Recognition, Barcelona, Spain, vol. 2*, pp. 1042–1045 (September 2000)
5. Lee, H.C., Gaensslen, R.E. (eds.): *Advances in Fingerprint Technology*. Elsevier, New York (1991)
6. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: FVC2000: Fingerprint Verification Competition. In: *15th IAPR International Conference on Pattern Recognition, Barcelona, Spain, September 3-7 (2000)*, <http://bias.csr.unibo.it/fvc2000/>
7. Pankanti, S., Prahakar, S., Jain, A.K.: On the individuality of fingerprints. *IEEE Trans. Pattern and Mach. Intell.*, 1010–1025 (2002)
8. Ratha, N., Chen, S., Jain, A.: An Adaptive flow orientation based feature extraction in fingerprint images. *Pattern Recognition*, 1657–1672 (1995)
9. Singh, R., Shah, U., Gupta, V.: Fingerprint Recognition, Student project, Department of Computer Science and Engineering, Indian Institute of Technology, Kanpur, India (November 2009)
10. John, V.C.: Fingerprint Based Authentication System, Correlation Based Classification Approach, MSITRT, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA
11. Bhanu, X.T., Yingqiang Lin, B.: Fingerprint classification based on learned features. Center for Res. in Intelligent Syst., Univ. of California, Riverside, CA (2005)

Modeling and Simulation of Respiratory Controller Using Labview

P. Srinivas¹, V. Naga Prudhvi Raj¹, and S. Rajesh²

¹ Associate Professor, Department of EIE, VR Siddhartha Engineering College,
Vijayawada, India
{paruchuris_99, nagaprudhvi}@yahoo.com

² Assistant Professor, Department of CSE, VR Siddhartha Engineering College,
Vijayawada, India

Abstract. The human respiratory system is a well-developed and complex system involving many different organs such as the nasal cavity, pharynx, trachea and the lungs. Though the actual physiological function of breathing begins only at birth, the development of the respiratory tract, the diaphragm and the lungs occur much earlier in the embryonic stage. Structural and functional changes continue from infancy to adulthood and into old age, as the respiratory system matures with age. Modeling of the respiratory system is helpful in finding out the diseases related to lungs. Engineering performed an effective support, bringing its namely in the construction of models and the simulation of the same. In this paper we develop an automatic controller for inhalation of oxygen through cylinders.

1 Introduction

For the construction of a system in the area of the medicine must be based on information in anatomy and physiology, then translate this knowledge in a similar system, which can be electrical, mechanical, hydraulic, chemical, etc, as the case and make the construction of the model, and be able to rely on a system of equations that can be validated in a simulation software. We have noted that the need for tools of modeling and simulation that will enable validate the body operation in normal and abnormally conditions. The engineers can be to design and develop mathematical models of human body that can be brought to a software and carry out the simulation that is validated by medical specialists [1][10].

The influence of technological changes can be detected in all aspects that concern the society due to the versatility and accessibility that offer. The systems can be designed to support the work done by the people, either in the industry, academic, medicine. In the studies medicine have been done profound, for several decades, to hear pathologies and power treated properly and have enjoyed the support of technological equipment that serves and works together with doctors[2].

2 Anatomy of Respiratory Tract

The respiratory system consists of an upper respiratory tract or via driving air formed by the nose, mouth, the pharynx, larynx and trachea; with each breath the upper tract leads the air into the interior and exterior of the ducts and structures that are the lower respiratory tract and the lungs. The main channels and structures in the lower respiratory tract are the trachea and, within the lungs, the bronchi, the bronchioles and the alveoli. In the depth of the lung, each bronchus is divided into bronchi secondary and tertiary, which continue branch in airways smaller than are called bronchioles. They end up in air sacs that are called alveoli. Refers to air introduced by the nose going through a series of passageways formed by the way nasal and is filtered by the of passageways formed by the way nasal and is filtered by the and then to the larynx, with a epiglottis, valve of soft tissues to seal the larynx during the act of swallowing so that prevents food will be introduced into the deep areas of respiratory tract. Under the larynx is the trachea, tubular structure that makes the transition between the tracts upper and lower its bottom is divided in two bronchi that embedded one in each lung, in turn bronchial tubes subdivided into multiple small ramifications, bronchioles and end with the formation of groups of small bags called alveoli. Each alveolus consists of a membrane fine and elastic serving as an interface between the breathing spaces of the lungs and blood that crosses the capillaries in the walls of the alveoli.

The oxygen content in the air and that reaches the alveoli presents dissemination through the membrane to reach the bloodstream, there is captured by the erythrocytes for distribution throughout the body. Simultaneously, the carbon dioxide that has been accumulated in the blood as a result of cellular metabolism presents dissemination in the opposite direction, toward the alveoli, to be expelled on expiration[2].

Two main types of lung disease, obstructive and restrictive, are related to changes in the following respiratory parameters: resistance and elastance. Restrictive lung diseases are caused either by an alteration in lung tissue, by disease of the chest wall, or by neuromuscular apparatus. A decrease is noted in the lungs ability to expand, or a decrease in the lungs' ability to transfer O_2 to the blood (or CO_2 out of the blood). In these conditions, the total lung volume and the transfer of oxygen from air to blood may be reduced. Restrictive disorders include sarcoidosis, interstitial pneumonitis, pulmonary fibrosis, and pneumonia. In obstructive lung conditions, airways are narrowed, usually causing an increase in the time that the lungs spend to be emptied. Obstructive lung disease can be caused by conditions such as emphysema, bronchitis, infection (which produces inflammation), and asthma; it also includes the common COPD [8][9].

In order to control the respiratory problems and ventilator for the patients we can develop an an automatic closed loop controller. This closed loop controller can be used to for sea divers and mountain trekkers for automatic opening and closing of the valve based on the oxygen available in the atmosphere. This method avoids manual control of the valve. The below figure [1] is the closed loop control system that can be designed. The respiratory output is given to the controller and based on the error the controller pumps the required amount of oxygen to the lungs. The closed loop control is shown in the following figure.

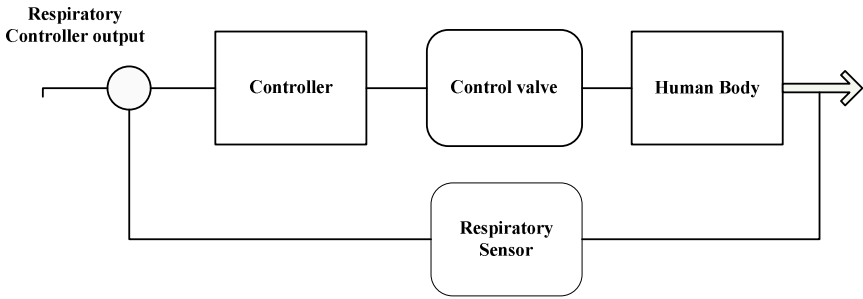


Fig. 1. Closed Loop Control

The controller is main in the closed loop control, so, for designing the control algorithm we first develop the model of the respiratory controller and simulate it. The simulation results are stored in the lookup table and based on that the controller error can be calculated. The data acquisition and control is easier in LABVIEW, hence the respiratory controller is simulated in LABVIEW. LABVIEW (short for Laboratory Virtual Instrumentation Engineering Workbench) is a platform and development environment for a visual programming language from National Instruments. In order to simulate, we need to model the respiratory system.

3 Modeling

Ventilation is highly sensitive to the partial pressure of CO_2 in arterial blood ($P_{a\text{CO}_2}$). A rise in $P_{a\text{CO}_2}$ by 1mm of Hg from its normal level of approximately 40mm Hg may increase the ventilator output by a third of its resting level. However, upon ascent to altitude or during inhalation of a gas mixture containing low O_2 content, there is an additional drive to breathe due to hypoxia. This hypoxia drive becomes noticeable when the partial pressure of O_2 in arterial blood, $P_{a\text{O}_2}$, drops below 70mm Hg. Since the metabolic consumption rate of O_2 and the metabolic elimination rate of CO_2 are relatively constant in the steady state, a higher level of ventilation would lead to an increase in $P_{a\text{O}_2}$ and a decrease in $P_{a\text{CO}_2}$, which in turn would lower ventilation. Therefore, the negative part of this negative feedback system is embedded in the gas exchange characteristics of the lungs. The ventilatory control system is divided into two components: the gas exchanging portion and the respiratory controller [4][5].

3.1 The Model for Gas Exchanger

Here we will restrict our attention only to gas exchange occurring in the lungs. The operating characteristics of the gas exchanger are obtained by deriving the mass balance equations for CO_2 and O_2 . First consider the CO_2 exchange only. We assume the metabolic CO_2 production rate to be V_{CO_2} , this is the rate at which CO_2 is delivered to the lungs from the blood that is perfusing the pulmonary circulation. In the steady state, this must equal the net flow of CO_2 exiting the lungs in gas phase. The latter is equal to the difference in volumetric fraction of CO_2 in the air entering (F_{ICO_2}) and leaving (F_{ACO_2}) the alveoli multiplied by the alveolar ventilation, V_A . The

4 Simulation

The model in LABVIEW is implemented and the input and output graphs are shown below. In LABVIEW we use simulation loop to use the threshold operator and integrator block which are present in control design and simulation module. In the front panel we represent the input graph that is a random sequence generator, output is graph plotted for input versus controller output[Figure 2,4] and graph plotted for partial pressures of CO_2 versus O_2 [Figure 3,5]. (8) Becomes progressively less valid as P_{aO_2} approaches the asymptotic value 38.6 as controller output becomes infinitely large. Here case (1) corresponds to the controller output at normal sea level. Case (2) corresponds to the controller output at an altitude of 8500 feet.

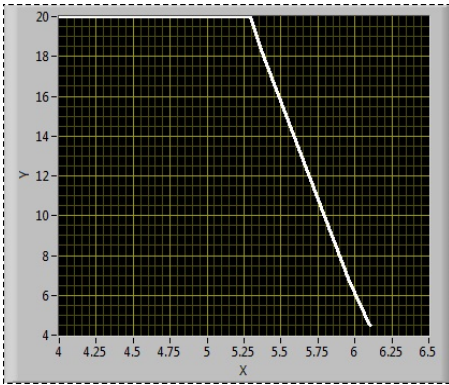


Fig. 2. V_{Ein} vs V_{Eout} in case (1)

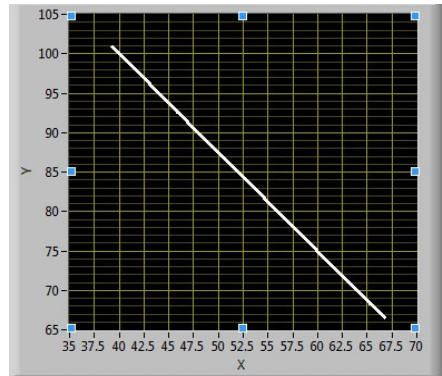


Fig. 3. P_{aO_2} vs P_{aCO_2} in case (1)

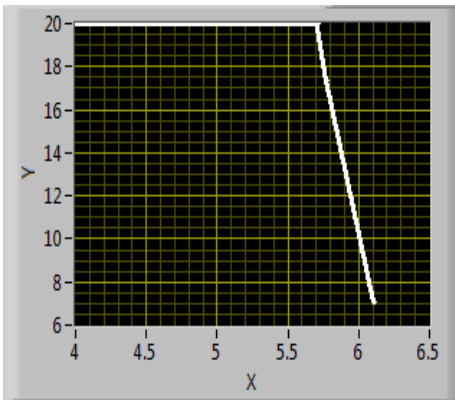


Fig. 4. V_{Ein} vs V_{Eout} in case (2)

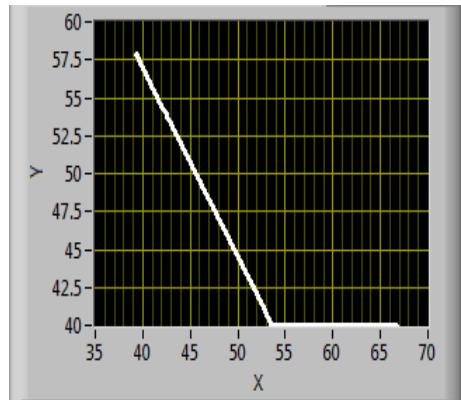


Fig. 5. P_{aO_2} vs P_{aCO_2} in case (2)

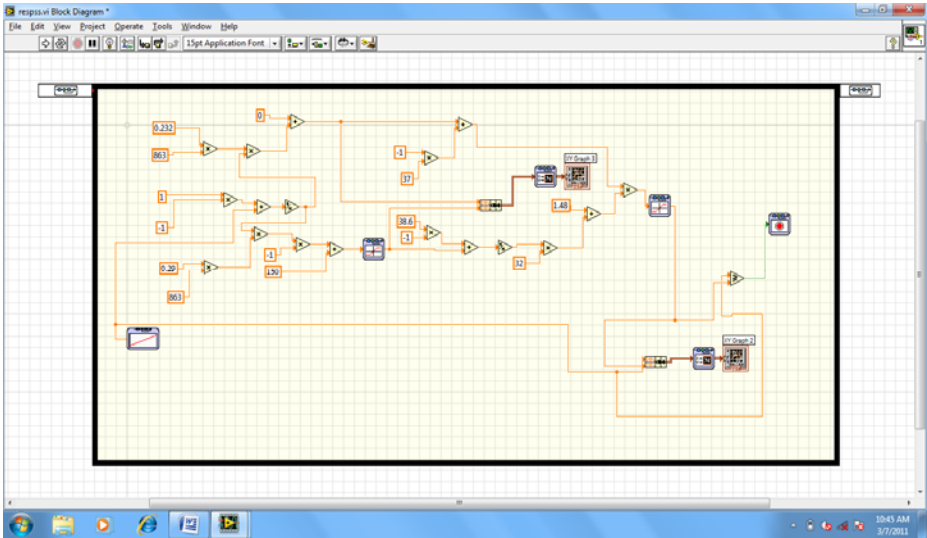


Fig. 6. Back panel

5 Observations

In case (1) we consider that if the person is at ground level. So the P_{IO_2} value is set equal to 150mm of Hg i.e. 21% of room air, while P_{ICO_2} is set equal to zero. Initially due low value of V_E , P_{AO_2} is 65 mm of Hg and P_{ACO_2} is 67 mm of Hg. As V_E increases, P_{ACO_2} decreases and P_{AO_2} increases. The simulation is terminated when $V_E=6$ L/min then $P_{ACO_2}=40$ mm of Hg and $P_{AO_2}=100$ mm of Hg. The values are tabulated below in table (1),(2)

Table 1. P_{AO_2} vs P_{ACO_2} in case 1

P_{AO_2} in mm of Hg	P_{ACO_2} in mm of Hg
67	65
60	70
50	80
40	100

Table 2. V_E in vs V_E out in case 1

V_E in	V_E out
5.1	20
5.4	15
5.8	10
6	6

In case (2) we consider that the person is at an altitude of 8500 feet. So we set P_{IO_2} value equal to 107 mm of Hg. while P_{ICO_2} is set equal to zero. Initially due low value of V_E , P_{AO_2} is 65 mm of Hg and P_{ACO_2} is 67 mm of Hg. As V_E increases, P_{ACO_2} decreases and P_{AO_2} increases. The simulation is terminated when $V_E=6.1$ L/min then $P_{ACO_2}=39$ mm of Hg and $P_{AO_2}=58.2$ mm of Hg. The values are tabulated below in table (3),(4).

Table 3. P_{AO_2} vs P_{ACO_2} in case 2

P_{AO_2} in mm of Hg	P_{ACO_2} in mm of Hg
40	65
40	60
45	52
58.2	39

Table 4. V_E in vs V_E out in case 2

V_E in	V_E out
5.1	20
5.4	13
5.8	9
6.1	6.1

6 Conclusions

From the above tables we can say that as the altitude increases the oxygen content decreases and we need to inhale more air than the required air. Here in this paper we developed a model for respiratory controller of human body which is used to calculate the steady state values of the output of controller. The values are entered in to look up tables and an controller can be developed to control the inhalation of oxygen. By implementing this control action we can reduce the wastage of the oxygen and we can avoid the manual control. The results show how the basic engineering tools help in the analysis of different systems, including biomedicine. Furthermore, the analysis of physiological systems can be reduced to simple electrical systems.

This can be further extended to medical application such as designing an automatic ventilator for patients. It also helps in various diagnosis related to respiratory system.

References

- [1] Cummingham, D.J.C.: Integrative aspects of regulation of breathing: A personal view. In: MTP International Review of Science: L Physiology Series One, vol. 2, pp. 303–369. Respiratory Physiology, University Park Press, Baltimore (1974)

- [2] Freedman, J., Lehmann, H., McCormick, R., Wachter, M.: Respiratory Physiology. Johns Hopkins School of Medicine, USA (1996)
- [3] Poon, C.S.: Respiratory Models and Controls. In: Bronzino, J.D. (ed.) Biomedical Engineering Handbook, 2nd edn., vol. 3. CRC Press, Boca Raton (1995)
- [4] Khoo, M.C.K., Kronauer, R.E., Strohl, K.P., Slutsky, A.S.: Factors inducing periodic breathing in humans: a general model. *J. Appl. Physiol.* 53, 644–659 (1982)
- [5] Diong, B., Nazeran, H., Nava, P., Goldman, M.: Modeling Human Respiratory Impedance. Comparing the Best Method with the Least Estimation Errors. *IEEE Engineering In Medicine And Biology Magazine* (January/February 2007)
- [6] Joseph, I.M.P., Butera, R.J.: A simple Model of Dynamic Interactions between Respiratory Centers. In: Proceedings of 27th Annual Conference, IEEE Engineering in Medicine and Biology, Shanghai, China, September 1-4 (2005)
- [7] Khoo, M.C.K.: Physiological Control Systems Analysis Simulation and Estimation. IEEE Press Series on Biomedical Engineering
- [8] Woo, T., Doing, B., Mansfield, L.: A Comparison of various Respiratory system models based on parameter estimates from Impulse oscillometry data. In: 26th Annual International Conference of the IEEE EMBS, San Francisco, CA, USA, September 1-5 (2004)
- [9] Koumandours, E., Santamaria, J.: The Comparison of work of breathing methodologies on a patient model. In: 23rd Annual EMBS International Conference, Isthambul, Turkey, October 25-28 (2001)
- [10] Guyton, A.C.: Textbook of Medical Physiology, 7th edn. Saunders (1986)
- [11] Mananas, M.A., Romero, S., Caminal, P.: A Comparative study of respiratory models in control of ventilation. In: Proceedings of Annual International Conference IEEE Engineering in Medicine and Biology, pp. 5491–5495 (2000)
- [12] Khoo, M.C., Krounauer, R.E.: Factors inducing periodic breathing in humans a General model. *Journal of Applied Physiology* 70(5), 2014–2024 (1991)
- [13] Otis, A.B., Fenn, W.O., Rahn, H.: Mechanics of breathing in man. *Journal of Applied Physiology* 15, 325–336 (1960)
- [14] Mananas, M.A.: Control system response of different respiratory models under ventilator stimuli and pathologies. In: Proceedings of 15th IFAC World Congress on Automatic Control, pp. 2317–2322 (2002)
- [15] Sparacino, M., Cobelli, C., Caumo, A.: Compartmental models of physiological systems. In: Medical Imaging Handbook, 2nd edn. CRC Press, Boca Raton (2000)
- [16] Roa, L., Ortega-Martinez, J.: Dynamic analysis of the respiratory system. In: 19th International Conference, IEEE EMBS, Chicago, October 30-November 2 (1997)
- [17] Bracio, B.R., Breitenecker, F.: Hybrid modeling and simulation of the Human respiratory system. In: 1st Joint BMES/EMBS Conference, Surviving Humanity Advancing Technology, Atlanta, USA, October 13-16 (1999)
- [18] Merit, J., Guha, R.K.: Simulation of lung respiration function using soft body model. In: 4th Burgeon Modeling Symposium on Computer Modeling and Simulation. IEEE, Los Alamitos (2010)
- [19] Tomlinson, S.P., Lo, J.: Time transient Gas Exchange in the respiratory system. *IEEE Engineering in Medicine and Biology Magazine* (September 1993)
- [20] Doing, B., Nazeran, H.: Modeling human respiratory impedance. *IEEE Engineering in Medicine and Biology Magazine* (January 2007)

A Comparative Study on Human Spermatozoa Images Classification with Artificial Neural Network Based on FOS, GLCM and Morphological Features

V.S. Abbiramy¹ and A.Tamilarasi²

¹ Velammal Engineering College, Chennai, India

² Kongu Engineering College, Perundurai, Erode

vml_rithi@yahoo.co.in,

drtamil@kongu.ac.in

Abstract. This paper aims to evaluate the accuracy of artificial neural network based classifiers using human spermatozoa images. Three different neural network based classifiers are used: Feed Forward Neural Network, Radial Basis Neural Network and Elman Back Propagation Neural Network. These three different classifiers were investigated to determine their ability to classify various categories of human spermatozoa images. The investigation was performed on the basis of the different feature vectors. The feature vector includes first order statistics (FOS), textural and morphological features. The extracted features are then used to train and test the artificial neural network. Experimental results are presented on a dataset of 91 images consisting of 71 abnormal images and 20 normal images. The radial basis network produced the highest classification accuracy of 60%, 75% and 70% when trained with FOS, Combined and Morphological features. When feed forward neural network is trained with GLCM features, a classification accuracy of 75% is achieved.

Keywords: Segmentation, Gray level Co-occurrence matrix, First Order Statistics, Artificial Neural Network, Image Processing, Classification, Spermatozoa.

1 Introduction

Over the last 10 years, attempts have been made to identify the correlation between spermatozoa morphology and its fertilizing capability [10]. Different staining techniques such as nigrosin-eosin stain, toluidine blue stain have been devised for examining sperm morphology. The stains have several drawbacks as its high cost in terms of time, specialized staff and the required equipments. The most preferred non staining technology for evaluating sperm morphology is Differential interference contrast microscopy [5].

Sperm Class Analyzer was used for the morphometric analysis of stallion sperm heads and midpieces. Harris Haematoxylin could be considered the most accurate staining method with the SCA [3]. The article discusses the Automated Sperm Morphology Analyzer (ASMA), a computer aided tool for the objective analysis of human sperm morphology [6]. Images of spermatozoa obtained with an optical phase-contrast microscope were segmented and feature vector for each head were computed. Learning vector quantization (LVQ) was applied to classify the feature vectors [7].

Texture descriptors of sperm images were extracted by means of Discrete Wavelet transform (DWT) and a back propagation neural was used as a classifier [8]. The fuzzy fusion morphology was used to express different shapes of human spermatozoa. Fuzzy operators were used to create the relationship between abnormal sperm morphology and normal sperm [9].

Human visual system is capable of recognizing and distinguishing textures with ease. However it appears to be a much more difficult task for the computers to perform the task of characterizing and distinguishing textures. In paper [11], the author has proposed an algorithmic model for automatic classification of flowers using KNN classifier. The proposed algorithmic model is based on textural features such as Gray level co-occurrence matrix and Gabor responses. Paper [12] presents a methodology for identification and classification of bulk sugary food objects. The texture features were extracted using gray level co-occurrence matrix method. The multilayer feed forward neural network was developed to classify bulk sugary food objects. The work [13] aims at selecting useful features in critical angles and distances by Gray Level Co-occurrence Matrix (GLCM). Features were classified using K means method into 2 classes of malignant and benign. Images obtained by Magnetic Resonance Imaging (MRI) of Iranian important export cultivar of pomegranate Malase-e-Torsh in [14] were analyzed by texture analysis to determine Gray Level Co-occurrence Matrix (GLCM) and Pixel Run-Length Matrix (PRLM) parameters. To classify pomegranate into different classes, discriminant analysis was conducted using cross-validation method and texture features.

Classification is a multistep process that consists of the following steps:

- Determine the target feature that will best describe the human spermatozoa.
- Create data set consisting of a set of attributes that captures the characteristics of spermatozoa.
- Normalize the data.
- Partition the data set into training and testing sets.
- Select an appropriate classification algorithm.
- Train the network using the training set.
- Test the network using the test data set.
- Assess the accuracy of the classifier.
- Finally apply the classifier on data to detect the feature.

2 System Flow of the Framework

The system flow of the processing framework is show in the figure 1. First the image database is setup and the corresponding class labels are found for each of the images. Images are then were preprocessed which involves cropping, resizing images, converting the images from RGB format Gray scale image and degraded image are restored using inverse filtering. Image is then segmented using adaptive threshold segmentation algorithm. The segmented image is given as input to the feature extraction process. During feature extraction phase, FOS, GLCM textural and morphological

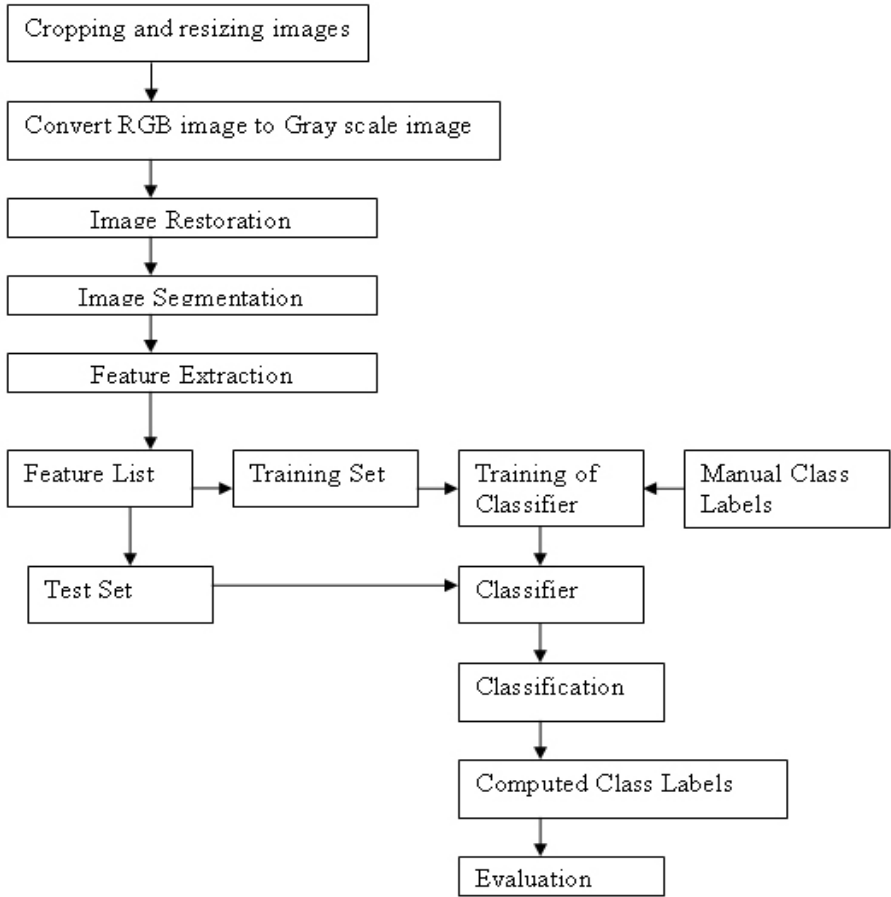


Fig. 1. System Flow

features are extracted for each of the image and saved in feature vector. Feature vector is then split into a training set and test set. The training set along with the corresponding class label is used to train the classifier. The trained classifier is then used to classify the images from the test set. During the evaluation phase, the computed results are compared with that of the actual results and the deviation is computed.

2.1 Preprocessing

The spermatozoon images used for this study are taken from World Health Organization laboratory manual [1]. These images are cropped and resized to a dimension of 86 x 100 pixels. The resized RGB images are then converted to gray scale images. Image restoration techniques are applied to improve the quality of the image and to remove noises from the images. The degraded image is restored using inverse filtering to extract a reliable GLCM, FOS and morphological features from the image.

2.2 Image Segmentation

Thresholding techniques produce segments having pixels with similar intensities. During the thresholding process, individual pixels in an image are marked as object pixels if their value is greater than some threshold value and as background pixels otherwise. Threshold based segmentation algorithm is used to segment the spermatozoon image to extract the FOS and textural features. Figure 2a shows the input image and Figure 2b shows the results of the segmentation using threshold based method.



Fig. 2. Results of Threshold based Segmentation (a) Input image (b) Segmented Image

To extract the morphological features, each and every spermatozoon image is subdivided into constituent parts such as head, nucleus and midpiece. Color based segmentation using K-Means clustering algorithm is used for subdividing the spermatozoon image. Figure 3 (a) shows the input image and Figure 3 (b), 3 (c), 3 (d) shows the head, midpiece and nucleus region of spermatozoa.

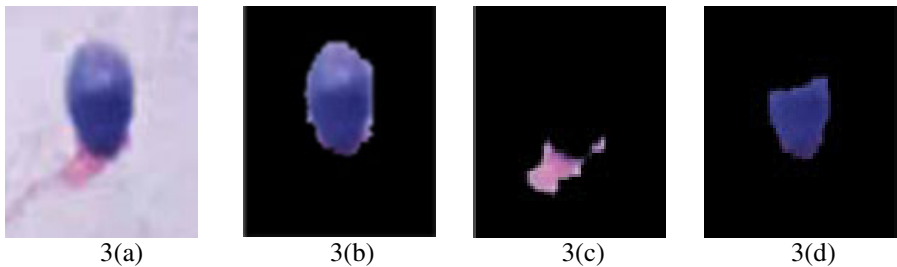


Fig. 3. Results of Color based Segmentation (a) Input image (b) Head (c) Midpiece (d) Nucleus

3 Feature Extraction

In this study three categories of features are extracted: First Order Statistics (FOS), Gray Level Co-occurrence Matrices and morphological features.

3.1 First Order Statistics Features:

FOS provides different statistical properties of the intensity histogram of an image. They depend only on individual pixel values and not on the interaction or

Table 1. FOS Features and their values

Image Type	Extracted Features			
	Mean	Deviation	skewness	kurtosis
Normal Image1	0.8517	0.3554	-1.9797	4.9192
Normal Image2	0.8366	0.3697	-1.8211	4.3163
Normal Image3	0.8271	0.3782	-1.7299	3.9925
Abnormal Image4	0.7906	0.4069	-1.4283	3.04
Abnormal Image5	0.8487	0.3583	-1.9464	4.7885

co-occurrence of neighboring pixel values. In this study, four first order textural features were calculated:

1. Mean value of gray levels
2. Standard Deviation of gray levels
3. Kurtosis: Kurtosis characterizes the relative peakedness or flatness of a distribution compared to the normal distribution and
4. Skewness: Skewness characterizes the degree of asymmetry of a distribution around its mean.

Table 1 shows the FOS features and values obtained for each of these features.

3.2 Grey Level Co-occurrence Matrix Features:

The GLCM [4] is a statistical method for extracting second order texture information from images. The GLCM characterizes the spatial distribution of gray levels in the

Table 2. GLCM Features and their values

S.No.	Features	Abnormal Image	Normal Image
1	Contrast	1.71418	2.77445
2	Correlation	0.04320	0.05409
3	Cluster Prominence	0.93987	0.97281
4	Cluster Shade	0.93987	0.97281
5	Dissimilarity	31.36181	165.17078
6	Energy	6.06725	23.11592
7	Entropy	0.04124	0.04626
8	Homogeneity	0.80763	0.74151
9	Maximum probability	0.54195	0.73096
10	Sum of squares	0.97971	0.97817
11	Sum average	0.97958	0.97765
12	Sum variance	0.89778	0.85952
13	Sum entropy	1.70014	2.75972
14	Difference variance	2.34695	2.68930
15	Inverse difference moment	4.76563	7.91695

segmented region. GLCM is computed by scanning the intensity of each pixel and its neighbor, defined by displacement d and angle θ [2]. In this study, four GLCMs corresponding to four different angles ($\theta=0^\circ, 45^\circ, 90^\circ$ and 135°) and one distance ($d=1$ pixel) were computed for each segmented image. Fifteen features were derived from each GLCM. The features extracted are: Contrast, Correlation, Cluster Prominence, Cluster Shade, Dissimilarity, Energy, Entropy, Homogeneity, Maximum Probability, Sum of Squares, Sum of Average, Sum of Variance, Sum of Entropy, Difference Variance and Inverse Difference moment. Four values were obtained for each feature corresponding to the four matrices. The second order textural features are obtained by calculating the mean (M) of these four values. Table 2 shows the GLCM features and values obtained for each of these features.

3.3 Combined Approach

It is a hybrid approach that combines first order textural features with that of the extracted GLCM features forming 19 features.

3.4 Morphological Features

In this stage, feature extraction is performed over the regions segmented in the previous stage. The features extracted represent the morphological parameters used in the assessment of human sperm morphology.

Table 3 shows the morphological features as in [15] and values obtained for each of these features.

Table 3. Morphological Features and their values

S.No.	Features	Abnormal Image	Normal Image
1	Head Area	6009	636
2	Head Perimeter	368	106.4264
3	Head Width	127.5424	35.6839
4	Head Length	113.1849	23.0324
5	Eccentricity	0.4609	0.7638
6	Orientation	5.6923	78.569
7	Midpiece Width	6.1604	23.9884
8	Midpiece Length	3.0742	13.2471
9	Nucleus	1413	352

4 Classification

In this paper, neural network is used for classification. The feature vectors returned by the feature extraction phase were first preprocessed before inputting them to the neural network. The extracted values are scaled to fall within the range of -1 and 1. The data was divided into two sets, one for training and the other for testing. Two layered Feed forward neural network, recurrent neural network and Radial basis neural network were then trained with the extracted FOS, GLCM and the combined features.

5 Experimental Results

Experiments were conducted with a dataset of 91 spermatozoa images consisting of 71 abnormal images and 20 normal images. Two third of the images were used for training and the remaining one third of the images were used for testing the network. The neural network is trained with the training set, whose classifier accuracy is estimated with the test set. The holdout method is repeated 7 times and overall classifier accuracy can be obtained by aggregating the individual holdout method values. Mean square error is the difference between the value predicted by the network and the actual values.

Two layered Feed forward neural network with learning function as traingdm and tansig as the activation function in hidden and output layer were employed. The number of neurons in the hidden layer was chosen as 5. For each training process, the sum square error performance goal is set to 0.0001 and maximum epoch is set to 300. Table 4 shows the classifier accuracy and mean square error of the feed forward neural network when trained with FOS, GLCM and the combined features.

Table 4. Classifier Accuracy of Feed Forward Neural Network

S.No	Training Rec	Test Rec	FOS		GLCM		COMBINED (FOS + GLCM)		Morphological Features	
			Accuracy	mse	Accuracy	mse	Accuracy	mse	Accuracy	mse
1	7	3	33	0.5756	67	0.3323	67	0.3291	67	0.8425
2	13	7	57	.1964	57	0.4252	57	0.4264	43	0.5248
3	20	10	60	0.3172	80	0.2006	70	0.2424	60	0.2880
4	27	13	84	0.1636	69	0.3009	62	0.3898	69	0.2824
5	33	17	24	0.3812	82	0.1765	82	0.1763	59	0.3425
6	40	20	50	0.1781	85	0.1264	85	0.1324	55	0.4403
7	47	23	34	0.1779	83	0.0910	83	0.1720	74	0.2560

Two layered Recurrent neural network with learning function as traingdm, tansig as the activation function in hidden layer , purelin as the activation function in the output layer were employed. The number of neurons in the hidden layer was set to 5. For each training process, the sum square error performance goal is set to 0.0001 and maximum epoch is set to 300. Table 5 shows the classifier accuracy and mean square error of the recurrent neural network.

Table 5. Classifier Accuracy of Recurrent Neural Network

S.No	Training Rec	Test Rec	FOS		GLCM		COMBINED (FOS + GLCM)		Morphological Features	
			Accuracy	mse	Accuracy	mse	Accuracy	mse	Accuracy	mse
1	7	3	66	2.0824	67	1.3047	67	0.5385	67	0.0852
2	13	7	43	0.3168	57	0.3397	57	0.9541	71	0.2020
3	20	10	50	0.8282	50	0.3531	50	0.6268	30	0.5165
4	27	13	85	0.1203	31	0.4703	46	0.3871	69	0.1427
5	33	17	71	0.9019	29	0.6841	29	1.0307	35	0.4692
6	40	20	0	0.2553	60	0.2040	90	0.2499	70	0.2942
7	47	23	91	0.1145	22	0.2435	61	0.1758	35	0.3036

A radial basis network is a network with two layers. It consists of a hidden layer of radial basis neurons and an output layer of linear neurons. Two parameters of the network are goal - mean squared error goal and spread of radial basis functions. Approximate choice of spread of RBF is required to fit a smooth function. Larger the spread, smoother the function approximation will be but too large spreads means a lot of neurons will be required to fit a large changing function. Also too small a spread means many neurons will be required to fit a smooth function and the network may not generalize well. Hence depending upon the mean absolute deviation of the input output pairs, the spread is approximatively chosen. Table 6 shows the classifier accuracy and mean square error of the radial basis neural network.

Table 6. Classifier Accuracy of Radial Basis Neural Network

S.No	Training Rec	Test Rec	FOS		GLCM		COMBINED (FOS + GLCM)		Morphological Features	
			Accuracy	mse	Accuracy	mse	Accuracy	mse	Accuracy	mse
1	7	3	33	4.5269	67	0.4616	67	0.3586	100	0.3532
2	13	7	86	2.1738	57	0.4572	71	0.4376	43	0.4831
3	20	10	60	7.3469	50	49.739	90	1.2164	60	0.3547
4	27	13	41	0.3876	62	2.2901	69	0.4283	85	0.1663
5	33	17	82	1.8862	59	7.5171	76	0.3332	71	0.1468
6	40	20	45	20.922	70	5.0869	80	0.4045	70	0.2405
7	47	23	70	5.7251	87	0.1673	74	0.2906	61	0.2643

Table 7 shows the performance comparison of feed forward, recurrent and radial basis neural network when trained with FOS, GLCM and the combined features (FOS + GLCM).

Table 7. Performance Comparison of the three types of Neural Networks

S.No.	Type of Neural Network	Classifier Accuracy of Features			
		FOS	GLCM	Combined	Morphological
1	Feed Forward	49	75	72	61
2	Recurrent	58	45	57	54
3	Radial Basis	60	65	75	70

6 Conclusion

In this paper, it has been shown that a two layered feed forward neural network seems to perform better (classification accuracy – 75%) when it is trained and tested with GLCM features. It correctly classified the samples that have large number of samples. Radial basis neural network could be employed successfully (classification accuracy - 60 %, 75% and 70%) as an image classifier with the FOS, combined technique (FOS + GLCM) and morphological features as the feature extractor.

It is suggested that the use of different distance measures at different angles could be used to improve the classification accuracy of similar look textures and rotated human spermatozoon images.

References

1. WHO: Laboratory Manual for the Examination and processing of human se-men: 5th edn., WHO (2010)
2. Karahaliou, A.N., et al.: Breast Cancer Diagnosis: Analyzing Texture of Tissue Surrounding Microcalcifications. *IEEE Transactions on Information Technology in Biomedicine* 12(6), 731–738 (2008)
3. Hidalgo, M., et al.: Effect of sample size and staining methods on stallion sperm morphology by the Sperm Class Analyzer. *Vet. Med.–Czech* 50(1), 24–32 (2005)
4. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural Features for image classification. *IEEE Transaction on System, Man and Cybermatics* 3(6), 610–621 (1973)
5. <http://www.vivo.colostate.edu/hbooks/pathphys/reprod/semeneval/morph.html>
6. Carrillo, H., et al.: A Computer Aided Tool for the Assessment of Human Sperm Morphology. In: *International Conference on Bioinformatics and Bioengineering, BIBE (2007)* 978-1-4244-1509-0
7. Alegre, E., et al.: Automatic classification of the acrosome status of boar spermatozoa using digital image processing and LVQ
8. Sanchez, L., et al.: Classification and quantification based on Image analysis for sperm samples with uncertain Damaged/ Intact Cell proportions. In: *International Conference on Image Analysis and Recognition, ICIAR (2008)*
9. Lu, J., Hu, Y.: *Soft Computing for Human Spermatozoa Morphology. Systemics, Cybernetics And Informatics* 3(3)
10. Kime, D.E., et al.: Computer-assisted sperm analysis (CASA) as a tool for monitoring sperm quality in fish. *Comparative Biochemistry and Physiology Part C* 130, 425–433 (2001)
11. Guru, D.S., et al.: Texture Features and KNN in Classification of Flower Images. In: *IJCA Special Issue on Recent Trends in Image Processing and Pattern Recognition, RTIPPR (2010)*
12. Anami, B.S., et al.: Texture based Identification and Classification of Bulk Sugary Food Objects. *ICGST-GVIP Journal* 9(4) (August 2009) ISSN: 1687-398X
13. Babaghorbani, P., et al.: Sonography Images for Breast Cancer Texture classification in Diagnosis of Malignant or Benign Tumors, 978-1-4244-4713-8/10/\$25.00 ©2010 IEEE
14. Khoshro, et al.: Classification of Pomegranate Fruit using Texture Analysis of MR Images. *Agricultural Engineering International: the CIGR Ejournal* XI (2009) (manuscript 1182)
15. Abbiramy, V.S., Shanthi, V.: Spermatozoa Segmentation and Morphological Parameter Analysis Based Detection of Teratozoospermia. *International Journal of Computer Applications* (0975 – 8887) 3(7) (June 2010)

Digital Video Watermarking Using Motion Detection and Singular Value Decomposition

Sanjana Sinha, Swarnali Pramanick, Ankul Jagatramka, Prajnat Bardhan,
Dipak K. Kole, and Aruna Chakraborty

Department of Computer Science Engineering,
St Thomas' College of Engineering and Technology,
Kolkata 700023, India

{sanjanasinha89, aruna.stcet, dipak.kole}@gmail.com

Abstract. Various digital watermarking techniques have been proposed in recent years for the purpose of protection of intellectual property rights of the creators of multimedia data, such as digital video. In order to achieve copyright protection of digital video we have adopted an imperceptible video watermarking approach based on Motion Detection and Singular Value Decomposition. Motion detection using temporal difference is used to identify the motion regions within a video frame where a watermark can be effectively embedded to ensure imperceptibility of the watermark. The unique properties of the singular values of an image, in this case the individual motion regions within the video frames, are utilized to enhance the robustness of watermarking and make it non-invertible. The resultant watermarked video is then tested for its robustness against attacks for watermark degradation and removal.

Keywords: Digital watermarking, copyright protection, motion detection, singular value decomposition.

1 Introduction

The popularity of digital video based applications is accompanied by the need for copyright protection to prevent illicit copying and distribution of digital video. Copyright protection inserts authentication data such as ownership information or logo in the digital media without affecting its perceptual quality. In case of any dispute, the authentication data is extracted from the media and can be used to prove the ownership of the digital media. As a method of copyright protection, digital watermarking [21][26] is the process by which a visible or invisible signal called a watermark or digital signature is inserted into a multimedia object such that watermark can be detected or extracted later to make an assertion about the object and resolve its rightful ownership. Technical challenges for providing a robust scheme for digital watermarking include various malicious attacks [11][12] like cropping, scaling, rotations and geometrical attacks. Different digital watermarking techniques [16] have been proposed over the past few years. Apparently any image watermarking technique [8][22] can be extended to watermarking video, but in reality video

watermarking techniques [5][6][7] need to meet difficult challenges [8] due to the properties of high correlation between successive video frames, the unbalance between the motion and motionless regions and real-time requirements during video broadcasting.

A variety of video watermarking schemes operate on uncompressed [27] and compressed video [6][19]. Most existing watermarking algorithms embed the watermark in the perceptually most insignificant components of the original video frames in order to guarantee the invisibility of the watermark. Video watermarking techniques are classified into spatial domain [27] and transform domain techniques [3]. In the spatial domain technique, the watermark is embedded in the source video by selecting pixel positions and replacing individual bits. Spatial domain techniques have lower complexity but are not as robust as compared to transform domain techniques. In the frequency domain, the values of chosen frequencies can be altered from the original. Since watermarks applied in the frequency domain will be dispersed over the entire spatial domain of the image upon inverse transformation, this method is not as susceptible to defeat by cropping as the spatial technique. In all frequency domain watermarking schemes, there is a conflict between robustness and transparency. If the watermark is embedded in perceptually most significant components, the scheme would be robust to attacks but the watermark may be difficult to hide. On the other hand, if the watermark is embedded in perceptually insignificant components, it would be easier to hide the watermark but the scheme may be less resilient to attacks. Hence there is a trade-off between invisibility and robustness.

In recent years, many watermarking schemes in the frequency domain have been developed using the Discrete Cosine Transform (DCT) [5], the Discrete Wavelet Transform (DWT) [4][9][14][20] and the Singular Value Decomposition (SVD) [1][13][15][28]. A significant amount of research has been carried out in image watermarking in the SVD domain. The SVD for square matrices was discovered independently by Beltrami in 1873 and Jordan in 1874, and extended to rectangular matrices by Eckart and Young in the 1930s. SVD is one of the most useful tools of linear algebra with several applications in image processing. In [18] an SVD based image watermarking technique is used which controls the strength of the watermark to be embedded. In [14] a blind video watermarking algorithm is proposed based on SVD. In [24] another SVD based algorithm is used to watermark video frames using a watermark whose strength can be controlled. The watermarks can be detected without the original video or any other information of the original singular values. In [9][10] a hybrid watermarking algorithm using SVD and DWT is used which embeds the watermark in the singular values of sub bands of DWT. With the hybrid approaches, the robustness of the scheme is increased. In [23] watermark bits are embedded in the Singular Value Decomposed video in a diagonal-wise fashion or in a block-wise fashion. In [29][30] improved digital watermarking algorithms based on Block-SVD are put forward, which have better robustness.

In this paper we propose a technique based on singular value decomposition of the motion regions identified within video frames through the process of motion detection.

The rest of this paper is organized as follows. Section 2 describes the proposed video watermarking technique. Section 3 presents the experimental results. The paper is concluded in Section 4.

2 Proposed Watermarking Technique

In this section, we propose a technique based on Singular Value Decomposition (SVD) for robust and imperceptible watermarking of digital video. The imperceptibility of watermarking is enhanced by embedding the watermark in the motion blocks of each video frame. The motion blocks are the two dimensional blocks of pixels having the highest motion activity within the video frames. Since the human visual system is less sensitive to high motion components in a video sequence, the watermark is embedded in regions of higher motion contained in motion blocks.

2.1 Singular Value Decomposition

Use of SVD in digital image processing has some advantages. Firstly, the size of the image matrices from SVD transformation can be $M \times N$ or $N \times N$. Secondly, singular values are less affected if general image processing is performed. Let an image be designated as a matrix A , denoted as $A \in F$, where F represents either the real number domain R or the complex number domain C .

The SVD A is defined as

$$A = USV^T$$

where $U \in F$ and $V \in F$ are unitary matrices, and $S \in F$ is a diagonal matrix. The diagonal entries of S (size $N \times N$) are called the singular values of A and are assumed to be arranged in decreasing order $\sigma_i > \sigma_{i+1}$. The columns of the U matrix ($M \times N$) are called the left singular vectors while the columns of the V matrix ($N \times N$) are called the right singular vectors of A . Each singular value specifies the luminance of an image layer while the corresponding pair of singular vectors specifies the geometry of the image layer. The singular values of an image have very good stability, i.e., when a small perturbation is added to an image, its singular values do not change significantly. Hence the watermark energy can be stored in the singular values of a video frame.

2.2 Motion Detection

Temporal difference [2] is an effective method for motion detection. The absolute difference between two consecutive video frames is calculated as D . I_n being the intensity of the n th frame, the difference image D_n is $(I_n - I_{n-1})$. Based on temporal difference, there is the block-based method to obtain the positions of the video to be watermarked where picture contents are moving fast. In the block-based method, the difference image D_n is divided into blocks. The sum of pixel values in each block is calculated and then the motion activity of each block is obtained. If the sum of pixel values in the blocks is greater, the motion activity of the block in the n th frame is higher.

In our work, we adopt the block-based method to select a series of blocks to form an area called motion region which has higher motion activity in each frame of the video sequence in order to embed the watermark.

2.3 Watermark Embedding Procedure

The high resolution video file is considered as a sequence of high resolution images called frames. The watermark embedding procedure consists of the following steps:

- 1) The video frames of size $N \times N$ are converted from *BGR* to *YCrCb* format. The *Y* component of each video frame is treated as a matrix.
- 2) The motion blocks of the video frame are used for embedding the watermark. For this purpose, the difference image D between two consecutive frames is calculated. D is subdivided into 4×4 sub-blocks. The subblocks having the highest additive sum of pixel intensities are grouped together to form a motion region R . The number of subblocks in motion region R are $n \times n / (4 \times 4)$. The subblocks are arranged to form a square matrix R .
- 3) The singular value decomposition of the R matrix of a video frame F is computed to obtain two orthogonal matrices U , V and one diagonal matrix S .

$$R \rightarrow USV^T . \quad (1)$$

- 4) We add a grayscale (single channel) watermark image W of size $n \times n$ to matrix S and perform SVD on the new matrix $S + \alpha W$ to get U_w , S_w and V_w .

$$S + \alpha W \rightarrow U_w S_w V_w^T , \quad (2)$$

where the positive constant ' α ' is the scale factor which controls the strength of the watermark and its values range from zero to one.

- 5) We then obtain watermarked motion region R_w by multiplying matrices U , S_w and V^T .

$$R_w \leftarrow US_w V^T . \quad (3)$$

- 6) The modified motion region R_w is placed in the original video frame to obtain the watermarked video frame A_w .

2.4 Watermark Extraction Procedure

Given matrices U_w , S , V_w and watermarked video frame A_w^* possibly corrupted by noise, filtering and other attacks, one can extract a possibly corrupted watermark image W^* by the following steps:

- 1) Divide the watermarked video into distinct frames.
- 2) Extract the luminance component of a frame (A_w^*).

- 3) The motion blocks of the video frame are again calculated to form a motion region R_w^* .
- 4) Apply SVD on R_w^* .

$$R_w^* \rightarrow U^* S_w^* V^{*T}. \quad (4)$$

- 5) Multiply U_w, S_w^* and V_w^T to obtain D^* .

$$D^* \leftarrow U_w S_w^* V_w^T. \quad (5)$$

- 6) Subtract S from D^* divided by the scale factor to obtain the extracted watermark.

$$W^* \leftarrow (1/\alpha)(D^* - R). \quad (6)$$

2.5 Method Analysis

In our method, we use block-based temporal difference to extract the motion region which has higher motion activity in each frame of video sequences than the rest of regions from the luminance layer of the video frame. It enhances the imperceptibility of the watermark in the watermarked video. The watermarking scheme is blind, robust and non-invertible. While many existing watermarking algorithms require the original image or video to extract the watermark, our method uses three matrices to extract the watermark and hence is a blind scheme. The imperceptibility of the watermarked video can be varied by changing the value of ' α '. Lower values of ' α ' result in lesser error between watermarked and original video frames but degrades the quality of the extracted watermark. Higher values of ' α ' will result in a good extracted watermark but the imperceptibility of watermarking will be affected.

The algorithm performs non-invertible watermarking [10][17] which means that it is computationally unfeasible for an attacker to find a faked video and a watermark such that the pair can result in the same watermarked video created by the real owner.

3 Experimental Results

The performance of the watermarking procedure is measured using well known metrics [25], such as *PSNR* (Peak-Signal-to-Noise-Ratio) and *NC* (Normalized Coefficient).

The Peak-Signal-To-Noise Ratio (*PSNR*) is used to measure the visual quality of watermarked and attacked frames and is defined as

$$PSNR = 10 \log_{10} (255^2 / MSE), \quad (7)$$

where *MSE* (mean squared error) between the original and distorted frames is defined as:

$$MSE = (1/mn) \sum_{i=1}^m \sum_{j=1}^n [I(i, j) - I'(i, j)]^2, \tag{8}$$

where m, n give the size of the frame and I and I' are the pixel values at location (i, j) of the original and the distorted frame respectively. A higher value of $PSNR$ indicates greater imperceptibility.

The normalized coefficient (NC) gives a measure of the robustness of watermarking and its peak value is 1.

$$NC = \frac{\sum_i \sum_j W(i, j) \cdot W'(i, j)}{\sqrt{\sum_i \sum_j W(i, j)^2} \sqrt{\sum_i \sum_j W'(i, j)^2}} \tag{9}$$

where W and W' represent the original and extracted watermark respectively. After extracting and refining the watermark, a similarity measurement of the extracted and the referenced watermarks is used for objective judgment of the extraction fidelity.

The proposed algorithm has been implemented on an AVI video clip of frame size 704×576 and frame rate 25 frames/sec and duration of 1 sec. The watermark used is a logo of size 255×255 . Both the video and the watermark are resized to an equal size before watermarking. Fig. 1 shows a single frame of the original video clip and Fig. 2 shows the watermarked video clip. Fig. 3(a) and Fig. 3(b) show the embedded and the extracted watermarks respectively. The watermarked video has a $PSNR$ of 48.0579.



Fig. 1. Original video clip



Fig. 2. Watermarked video clip



(a)



(b)

Fig. 3. (a) Embedded Watermark (b) Extracted Watermark

The watermarked video frames are subject to a number of geometric and filtering attacks. Fig. 4 shows the watermarked video frame subjected to rotation by an angle of -5 degrees. Fig. 5 shows the resultant video frame after addition of Gaussian noise (factor 2.5). Fig. 6 shows the frame after median filtering (3×3 box filter) and Fig. 7 shows it after addition of Poisson noise (factor 0.5). Fig. 8 shows the video frame after contrast adjustment (factor 30). Fig. 9 shows the resultant frame after sharpening filter (factor 50) is applied to the watermarked video frame.

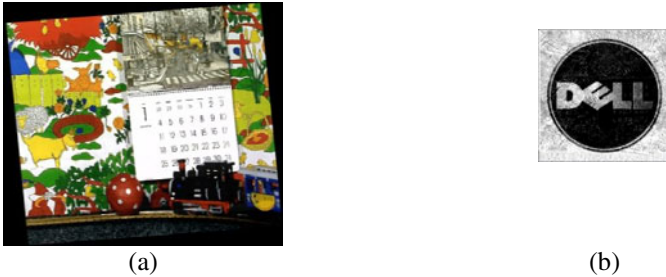


Fig. 4. (a) Watermarked frame after rotation by -5 degrees (b) Extracted watermark

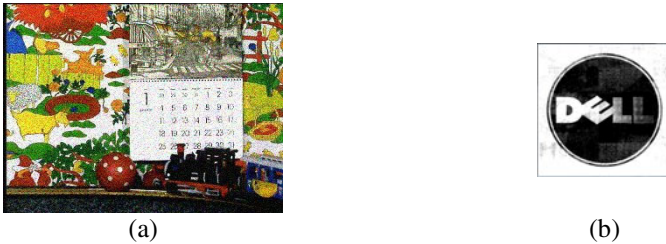


Fig. 5. (a) Watermarked frame after addition of Gaussian noise (factor 2.5) (b) Extracted watermark

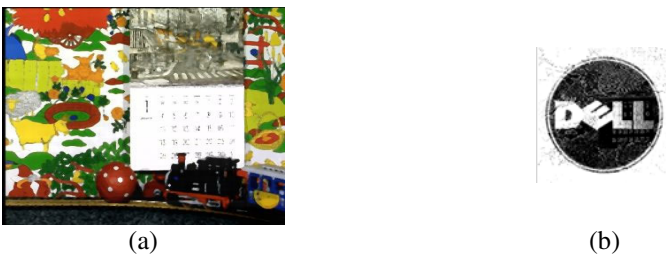


Fig. 6. (a) Watermarked frame after Median filtering (b) Extracted watermark

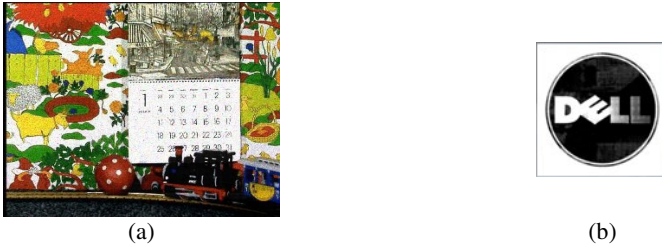


Fig. 7. (a) Watermarked frame after addition of Poisson Noise (b) Extracted watermark

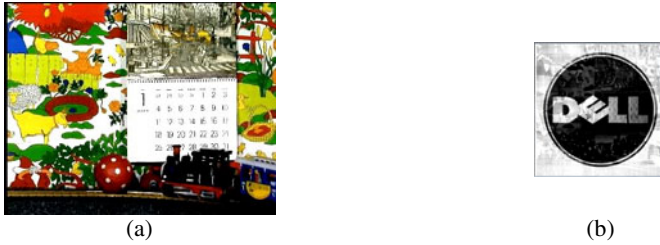


Fig. 8. (a) Watermarked frame after Contrast Adjustment (b) Extracted watermark



Fig. 9. (a) Watermarked frame after application of Sharpening filter (b) Extracted watermark

The following table gives the values of PSNR and NC corresponding to each of the attacks listed below.

Table 1. Test Results for Attacks on Watermarking

Attack name	PSNR(dB)	NC
Rotation	26.5465	0.8456
Resize	35.9926	0.9017
Cropping	27.0249	0.9238
Gaussian Noise	33.6837	0.9455
Median Filtering	37.3004	0.9316
Poisson Noise	39.6479	0.9738
Contrast Adjustment	34.8666	0.9714
Sharpening filter	34.9161	0.9011

4 Conclusion

The proposed watermarking technique is a robust and imperceptible scheme which embeds a digital image watermark into the frames of video in the uncompressed domain by utilizing the correlation between successive frames. The SVD transform technique has certain distinguished properties that make the algorithm robust against frame distortions. Since the human visual system is less sensitive to high motion components in a video sequence, embedding the watermark in regions of higher motion contained in motion blocks enhances the imperceptibility of the watermarking scheme. While this algorithm can be applied to individual frames of the video, a more efficient technique will be to use time-variant watermarks for video scenes using the concept of scene change analysis.

As a future work we propose a technique to combine the benefits of our algorithm with other transform domain techniques to create a hybrid algorithm that will enhance the security of the copyrighted video and make it feasible for real time implementation.

References

1. Abdallah, H.A., Hadhoud, M.M., Shaalan, A.A.: SVD-based watermarking scheme in complex wavelet domain for color video. In: International Conference on Computer Engineering & Systems, pp. 455–460 (2009)
2. Anderson, C., Burt, P., Van der Wal, G.: Change detection and tracking using pyramid transformation techniques. In: Proc. SPIE-Intelligent Robots and Computer Vision, vol. 579, pp. 72–78 (1985)
3. Baisa, L.G., Manthalkar, R.R.: An Overview of Transform Domain Robust Image Watermarking Algorithms. *Journal of Emerging Trends in Computing and Information Sciences* 2, 37–42 (2010)
4. Chetan, K.R., Raghavendra, K.: DWT Based Blind Digital Video Watermarking Scheme for Video Authentication. *International Journal of Computer Applications* (10), article 4 (2010)
5. Deshpande, N., Rajurkar, A., Manthalkar, R.: Review of Robust Video Watermarking Algorithms. *IJCSIS* 7(3), 237–246 (2010)
6. Ding, Y., Zheng, X., Zhao, Y., Liu, G.: A Video Watermarking Algorithm Resistant to Copy Attack. In: Third International Symposium on Electronic Commerce and Security (2010)
7. Doërr, G., Dugelay, J.L.: A guide tour of video watermarking. *Signal Processing: Image Commun.* 18(4), 263–282 (2003)
8. Doërr, G., Dugelay, J.L.: Security Pitfalls of Frame-by-Frame Approaches to Video Watermarking. *IEEE Transactions on Signal Processing* 52, 2955–2964 (2004)
9. Ganic, E., Eskicioglu, A.M.: Robust DWT-SVD Domain Image Watermarking: Embedding Data in All Frequencies. In: Proceedings of the Workshop on Multimedia and Security (2004)
10. Haj, A.A., Manasrah, T.: Non-Invertible Copyright Protection of Digital Images Using DWT and SVD. In: 2nd International Conference on Digital Information Management, ICDIM 2007, vol. 1, pp. 448–453 (2007)
11. Jayamalar, T., Radha, V.: Survey on Digital Video Watermarking Techniques and Attacks on Watermarks. *International Journal of Engineering Science and Technology* 2(12), 6963–6967 (2010)

12. Kalker, T.: Considerations on watermarking security. In: Proceedings IEEE Workshop Multimedia Signal Processing, pp. 201–206 (2001)
13. Kaufman, J., Celenk, M.: Digital Video Watermarking using Singular Value Decomposition and 2D Principal Component Analysis. In: IEEE International Conference on Image Processing, pp. 2561–2564 (2007)
14. Khatib, T., Haj, A., Rajab, L., Mohammed, H.: A Robust Video Watermarking Algorithm. *Journal of Computer Science* 4, 910–915 (2008)
15. Kong, W., Yang, B., Wu, D., Niu, X.: SVD Based Blind Video Watermarking Algorithm. In: First International Conference on Innovative Computing, Information and Control, ICIC 2006, pp. 265–268 (2006)
16. Langelaar, G., Setyawan, I., Lagendijk, R.: Watermarking Digital Image and Video Data: A State of - Art Overview. *IEEE Signal Processing Magazine*, 20–46 (2000)
17. Li, Q., Chang, E.: On the Possibility of Non-Invertible Watermarking Schemes. In: Proceedings of the 6th International Workshop on Information Hiding, Canada, pp. 13–24 (2004)
18. Liu, R., Tan, T.: A SVD-Based Watermarking Scheme For Protecting Rightful Ownership. *IEEE Transactions on Multimedia* 4, 121–128 (2009)
19. Meng, J., Chang, S.: Embedding Visible Video Watermarks in the Compressed Domain. In: Proceedings of International Conference on Image Processing, ICIP 1998, vol. 1, pp. 474–477 (1998)
20. Patel, H., Patoliya, J., Panchal, P., Patel, R.N.: Digital Robust Video Watermarking using 4-level DWT. *International Journal of Advanced Engineering Technology* 1, 101–113 (2010)
21. Podilchuk, C.I., Delp, E.J.: Digital watermarking: algorithms and applications. *IEEE Signal Processing Magazine* 18, 33–46 (2001)
22. Potdar, V., Han, S., Chang, E.: A Survey of Digital Image Watermarking Techniques. In: Proceedings of the IEEE International Conference on Industrial Informatics, pp. 709–716 (2005)
23. Rajab, L., Khatib, T., Haj, A.: Video Watermarking Algorithms Using the SVD Transform. *European Journal of Scientific Research* 30(3), 389–401 (2009) ISSN 1450-216X
24. Sinha, S., Pramanick, P., Jagatramka, A., Bardhan, P.: Digital Video Watermarking using Singular Value Decomposition. In: Proceedings on IEEE EDS Student Paper Conference, pp. 53–56 (2011)
25. Thakur, M.K., Saxena, V., Gupta, J.P.: A Performance Analysis of Objective Video Quality Metrics for Digital Video Watermarking. In: 3rd IEEE International Conference Computer Science and Information Technology (ICCSIT), vol. 4, pp. 12–17 (2010)
26. Tirkel, A.Z., Rankin, G.A., Van Schyndel, R.M., Ho, W.J., Mee, N.R.A., Osborne, C.F.: Electronic Water Mark. In: DICTA 1993, pp. 666–673. Macquarie University (1993)
27. Tokari, T., Kanoczi, T., Levicky, D.: Digital Watermarking of uncompressed video in spatial domain. In: 19th International Conference on Radioelektronika, pp. 319–322 (2009)
28. Wu, Y.: On the Security of an SVD-Based Ownership Watermarking. *IEEE Transactions on Multimedia* 7(4) (2005)
29. Zhou, Z., Tang, B., Liu, X.: A Block-SVD Based Image Watermarking Method. In: Proceedings of the 6th World Congress on Intelligent Control and Automation (2006)
30. Zhu, X., Zhao, J., Xu, H.: A Digital Watermarking Algorithm and Implementation Based on Improved SVD. In: 18th International Conference on Pattern Recognition, ICPR, vol. 3, pp. 651–656 (2006)

Mining Association Rules Using Hash-Index Technique

R.B. Geeta¹, Omkar Mamillapalli², Shasikumar G. Totad³,
and P.V.G.D. Prasad Reddy⁴

¹Associate Professor, Department of CSIT, GMRIT, Rajam, (A.P), India

²Department of CSE, GMR Information Technology, Rajam, (A.P), India

³Professor & HOD, Department of CSE, GMRIT, Rajam, (A.P), India

⁴Professor, Department of CS & SE, Andhra University, Visakhapatnam, (A.P), India

{omkar526, prasadreddy}@gmail.com,

geetatotad@yahoo.co.in, skumartotad@yahoo.com

Abstract. This paper presents the hash index table structure, a general and dense structure which provides web page set extraction from Log File of server. This hash table provides information about the original database. Since no constraint is enforced during the index creation phase, Web Page set mining (WPs-Mine) provides a complete representation of the original database. This approach works well for both sparse and dense data distributions. The proposed web page set index is linearly scalable for large data sets. Web page set mining supported by hash table index shows the performance always comparable with and often better than algorithms accessing data on flat files. Incremental update is feasible without re accessing the original transactional database.

Keywords: Data mining, Web Pages set extraction, Hash indexed tree.

1 Introduction

Association rule mining has been investigated by many researchers and practitioners for many years [1], [2], [3], [4], [5]. Association rules are used to identify correlations among web pages in a transactional database D . Association refers to the data mining task of uncovering relationships among data. An Association Rule is a model that identifies specific types of data associations. These are used to assist retail store management, marketing and inventory control. Association rule mining deals with finding frequent patterns, associations, correlations.

Each transaction in D is a set of web pages. Association rules are usually represented in the form $A \rightarrow B$, where A and B are web page sets, i.e., set of web pages. Web page sets are characterized by their frequency of occurrence in D , which is called support. Research activity usually focuses on defining efficient algorithms for web page set extraction, which represents the most computationally intensive knowledge extraction task in association rule mining [6]. In this paper, we propose a similar approach to support data mining queries. The WebPages-Mine (WPs-Mine) index is a novel data structure that provides a compact and complete representation of transactional data supporting efficient item set extraction from a relational DBMS. The following Web Pages data set shows 13 hypertext transfer protocol (http) transactions requests in given session threshold.

Table 1. Example Web Pages data set

TID	WebPagesID	TID	WebpagesID	TID	WebPagesID
1	b,h,e,p,v,d,g	6	a, r, n, u ,i, b,s	11	r,e,h,b,a
2	m,h,n,d,b,e	7	b, g, h, d, e,p	12	z,i,a,n,r,b
3	l,f,o,h,e,c,p	8	a, i, b	13	b,d,h,p,e
4	a,w,e,k,h,j	9	f, e ,i, c ,h, p		
5	d,b,e,h,n	10	h, a, e, b, r ,t		

2 The WPs-Mine Index

Whenever user requests for any hypertext transfer protocol (http) transaction, the details of request is entered into the Log File of the server. The log file entry contains various fields like IP address, time at which request is made, status code, number of bytes transferred and which page is requested. The web pages information collected in the log file is stored in the form of database. This data is stored in the form of relational model, as a relation R. Assuming some session threshold the frequency of each webpage is counted and stored in dataset as shown in table 1.

2.1 WPs-Hash Indexed Tree Structure

The structure of the WPs-Mine index is characterized by two components: the Web Page Set-Tree and the Web Pages-Hash table tree. The two components provide two levels of indexing. The Web Pages set-Tree (WPs-Tree) is a prefix-tree which represents relation R by means of a brief and compact structure. The hash table of 26 buckets [A-Z] is created. Each bucket stores the information about the support of each web page in a assumed threshold. Each bucket holds the physical location of each web page in the website. Linked list with various nodes is attached for each bucket which holds the addresses of different web pages. The WebPages-Hash index (WPs-H index) table structure allows reading selected WPs-Tree portions during the extraction task. For each item, it stores the physical locations of all item occurrences in the WPs-Tree.

2.1.1 WPs-Tree

The Web Pages set-Tree (WPs-Tree) is a prefix-tree which represents relation R by means of a short and compact structure. Implementation of the WPs-Tree is based on the FP-tree data structure, which is very effective in providing a compact and lossless representation of relation R as shown in Fig.1.

2.1.2 WPs-Hash-Indexed tree

The WPs-Hash-tree is a Hash table with tree structure which allows access of selected WPs-Tree portions during the extraction task. For each web page in the given website, it stores the physical locations of all web page occurrences in the Web Pages set Tree.

Fig. 2 show The WPs-Mine Hash indexed tree allows selectively accessing the WPs-Tree blocks during the extraction process. It is based on a Hash indexed Tree structure. For each item i in relation R, there is one entry in the WPs-Mine Hash indexed tree.

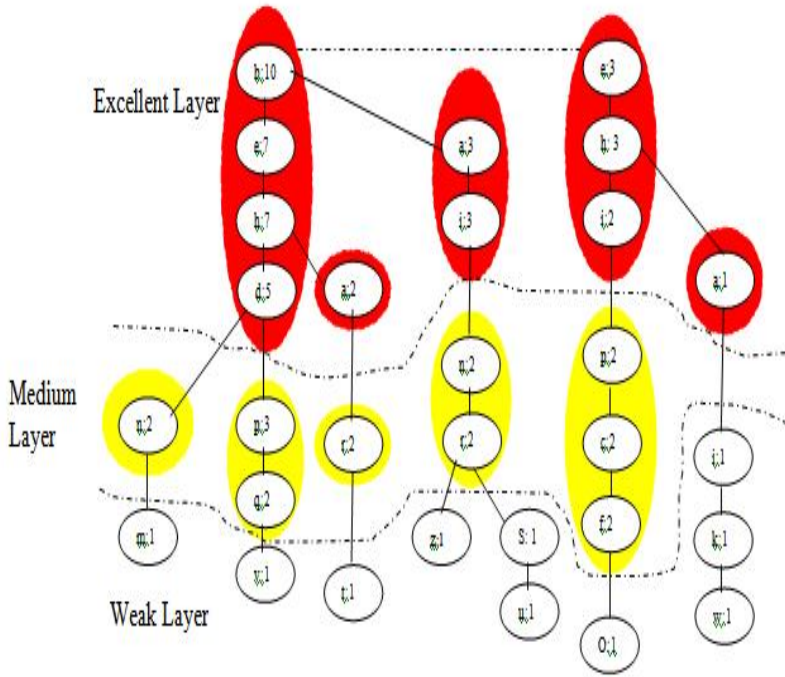


Fig. 1. WPs-Mine index for the example dataset WPs-Tree

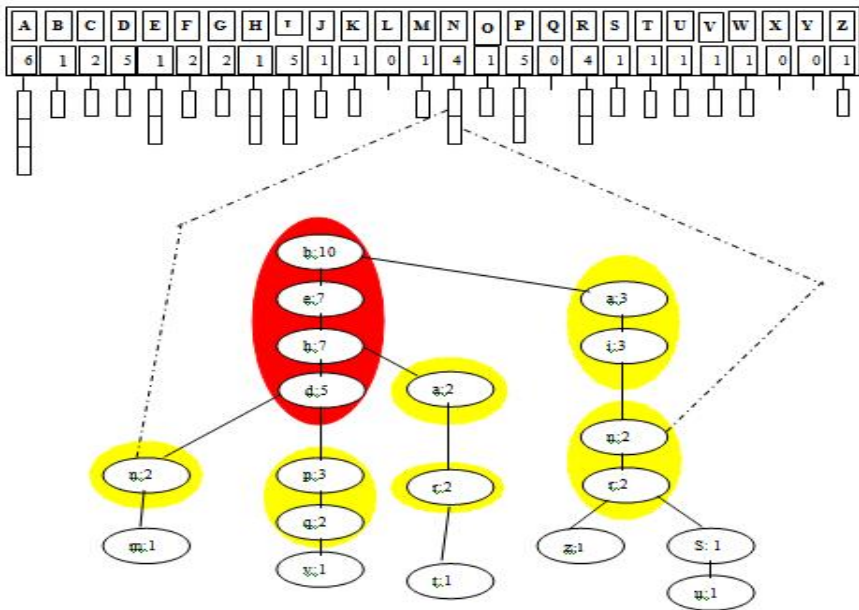


Fig. 2. WPs-Mine Hash indexed tree for the example dataset WPs tree

2.2 WPs-Mine Data Access Methods

Three data access methods are devised to load from the WPs-Mine index the following projections of the original database: 1) Frequent WebPages-Tree to support projection-based algorithms (e.g., FP-growth [7]). 2) Support-based projection, to support level based (e.g., APRIORI [6]), and array-based (e.g., LCM v.2 [8]) algorithms. 3) web pages-based projection, to load all transactions where a specific webpage occurs, enabling constraint enforcement during the extraction process. The three access methods are described in the following sections.

2.2.1 Construction of Frequent WebPages-Tree

From the relation R, the frequency of each web page is counted for a given session threshold time. The web pages are sorted in order based on its frequency but preceding in lexicographical order. In the WPs-Tree web pages are sorted by descending support lexicographical order as represented by WPs-Tree. This is represented as a prefix tree.

In the example data set, item p appears in two nodes, i.e., [p:3] and [p:2]. The access method reads two prefix paths for p, i.e., [p : 3 ->d :5 -> h : 7 !->e : 7 -> b : 10]and [p : 2 !->i : 2 !->h : 3 ->e:3] Each sub path is normalized to p node support. For example, the first prefix path, once normalized to [p:3], is [p : 3 ->d :3 -> h : 3 -> e : 3 !->b : 3]

2.2.2 Support-Based division of WPs-Tree

The support-based projection of relation R contains all transactions in R intersected with the web pages which are frequent with respect to a given support threshold (Min Sup). The WPs-Tree paths completely represent the transactions. Web pages are sorted by decreasing support along the paths.

Starting from a root node, the WPs-Tree is visited depth-first by following the node child pointer. The visit ends when a node with an Un-frequent item or a node with no children is reached.

The WPs-Tree is partitioned into three layers based on given minimum support threshold. The web pages whose support is greater than or equal to given minimum threshold is considered to be belonging to Excellent layer. The web pages whose support is greater than 1 and less than given minimum support belong to weak layer as shown in Fig. 1.

2.2.3 WPs-Hash-Table Tree

Log File of server contains information about how many visitors visited various web pages of web site. Given a session threshold, the frequency of each web page is counted and each page's count is stored in hash table. The hash table of 26 buckets [A-Z] is created. Each bucket stores the information about the frequency of each web page. Each bucket holds the physical location of each web page occurrences in the WPs-tree. Linked list with various nodes is attached for each bucket which holds the addresses of the occurrences of web pages in the WPs-Tree. The WebPages-Hash index (WPs-H index) table structure allows reading selected WPs-Tree portions during the extraction task. For each web page, it stores the physical locations of all page occurrences in the WPs-Tree.

2.3 WPs-Mine Physical Organization

The physical organization of the WPs-Mine index is designed to minimize the cost of reading the data needed for the current extraction process. However, fetching a given record requires loading the entire disk block where the record is stored. The WPs-Tree physical organization is based on the following correlation types:

i. **Intra transaction correlation:** Web Pages appearing in a same transaction are thus intrinsically correlated. To minimize the number of read blocks, each WPs-Tree path should be partitioned into a block.

ii. **Inter transaction correlation:** In some transactions, set of web pages accessed may be same and some other pages accessed may be different, so block can be formed for common web pages accesses and separate block can be made for remaining web pages access.

2.3.1 WPS-Tree Layers

The WPS-Tree is partitioned in three layers based on the node access frequency during the extraction processes. 1) the node level in the WPs-Tree, i.e., its distance from the root, 2) the number of paths including it, represented by the node support, and 3) the support of its item.. The three layers are shown in Fig. 2a for the example WPs-Tree.

Excellent layer: This layer includes web pages that are very frequently accessed during the mining process. These nodes are located in the upper levels of the WPs-Tree. These web pages are most important pages as these are frequently accessed.

Medium Layer: This layer includes nodes that are quite frequently accessed during the mining process. This layer contains web pages which are frequently accessed during web site visits.

Weak layer: This layer includes the nodes corresponding to rather low support items, which are rarely accessed during the mining process. The web pages in this layer must be paid more attention to modify the content as these web pages are rarely accessed by web users.

3 Web Page Set Mining

Web Page set mining has two sequential steps: 1) the needed index data is stored and 2) web page set extraction takes place on stored data.

3.1 Frequent Web Pages Set Extraction

This section describes how frequent web pages set extraction takes place on the WPs-Mine index. We present two approaches, denoted as FP-based and LCM-based algorithms, which are an adaptation of the FP-Growth algorithm [3] and LCM v.2 algorithm [6], respectively.

3.1.1 FP-Based Algorithm

The FP-growth algorithm stores the data in a prefix-tree structure called FP-tree. First, it computes web page support. Then, for each transaction, it stores in the FP-tree its subset including frequent web pages. Web pages are considered one by one. For each web page, extraction takes place on the frequent-web page database, which is generated from the original FP-tree and represented in a FP-tree based structure.

3.1.2 LCM-Based Algorithm

The LCM v.2 algorithm loads in memory the support-based original database. First, it reads the transactions to count item support. Then, for each transaction, it loads the subset including frequent web pages. Data are represented in memory by means of an array-based data structure, on which the extraction takes place.

3.2 Enforcing Constraints

Constraint specification allows the (human) analyst to better focus on interesting web page sets for the considered analysis task. Constraints have been classified as ant monotonic, monotonic, succinct, and convertible.

The approach in can be easily extended to deal with multiple constraints The WPs-Mine index provides a flexible information storage structure which can easily support also this extension.

Table 2. Data Set Characteristics and Corresponding Indices

Dataset	Transactions	Dataset Items	AvtrSz	Size(KB)	I-Tree (KB)	IMine I-Btree (KB)	Time(sec)
CONNECT	67,557	129	43	25,527	22,527	4,211	11.05
PUMSB	98,092	2,144	37,01	35,829	57,932	10,789	34,47
KOSARAK	1,017,029	41,244	7.9	85,435	312,647	58,401	893.81
T101200P20D2M	2,000,000	86,329	20.07	544,326	233,872	104,605	666.5
T151100P20C1D5M	5,000,000	45,666	22	1,476,523	1,464,144	277,029	3,736.7
T201100P15CID7M	7,000,000	39,141	22	2,075,478	6,758,896	944,450	8350.72

4 Experimental Results

We validated our approach by means of a large set of experiments addressing the following issues:

1. Performance of the WPs-Mine index creation, in terms of both creation time and index size,
2. Performance of frequent web pages set extraction, in terms of execution time, memory usage, and I/O Access time,

We ran the experiments for both dense and sparse data distributions. We report experiments on six representatives data sets whose characteristics (i.e., transaction

and item cardinality, average transaction size (AvgTrSz), and data set size) are in Table 1. Connect and Pumsb [8] are dense and medium-size data sets. Kosarak [8] is a Large and sparse data set including click-stream data. T10I200P20D2M is a dense and large synthetic data set, while T15I100P20C1D5M and T20I100P15C1D7M are quite sparse and large synthetic data sets.

4.1 Index Creation and Structure:

Table 2 reports both WPs-Tree and WPs-Hash index table tree size for the six data sets. The overall WPs-Mine index size is obtained by summing both contributions. The WPs-Mine indices have been created with the default value $K_{avg} \frac{1}{4} 1:2$. Furthermore, the Connect, Pumsb, Kosarak, and T10I200P20D2M data sets have been created with $K_{Sup} \frac{1}{4} 0$, while large synthetic data sets with $K_{Sup} \frac{1}{4} 0:05$. In dense data sets (e.g., T10I200P20D2M) where data are highly correlated, the WPs-Tree structure is more compact. In sparse data sets (e.g., Kosarak), where data are weakly correlated, data compression is low and storing the WPs-Tree requires more disk blocks. For example, with respect to the original database, the WPs-Tree for T10I200P20D2M has a lower size, for T15I100P20C1D5M is quite equivalent, and for Kosarak and T20I100P15C1D7M is larger.⁴ The WPs-Hash index tree contains pointers to all nodes in the WPs-Tree.

Table 1 also shows the index creation time, which is mainly due to path correlation analysis and storage of the index paths on disk. The first factor depends on the number of WPs-Tree paths.

4.2 Frequent Web Pages Set Extraction Performance

The WPs-Mine structure is independent of the extraction algorithm. To validate its generality, we compared the FP-based and LCM-based algorithms with three very effective state-of-the-art algorithms accessing data on flat file.

Fig. 3 compares the FP-based algorithm with the FP-growth algorithms [3] on flat file, all characterized by a similar extraction approach. For real data sets (Connect, Pumsb, and Kosarak), differences in CPU time between the FP-based and the Prefix-Tree algorithms are not visible for high supports, while for low supports the FP-based approach always performs better than Prefix-Tree. The FP-based algorithm, albeit implemented into a relational DBMS, more memory space is available for the extraction task. This effect is particularly relevant for low supports, because representing in memory a large portion of the data set may significantly reduce the space for the extraction task, hence causing more memory swaps.

As shown in Fig. 4, LCM-based approach provides an extraction time comparable to LCM v.2 on flat file. For large data sets, it performs better than LCM v.2. Since WPs-tree paths compactly represent the transactions reading the needed data through the index requires a lower number of I/O operations with respect to accessing the flat Tree, This benefit increases when the data set is larger and more correlated.

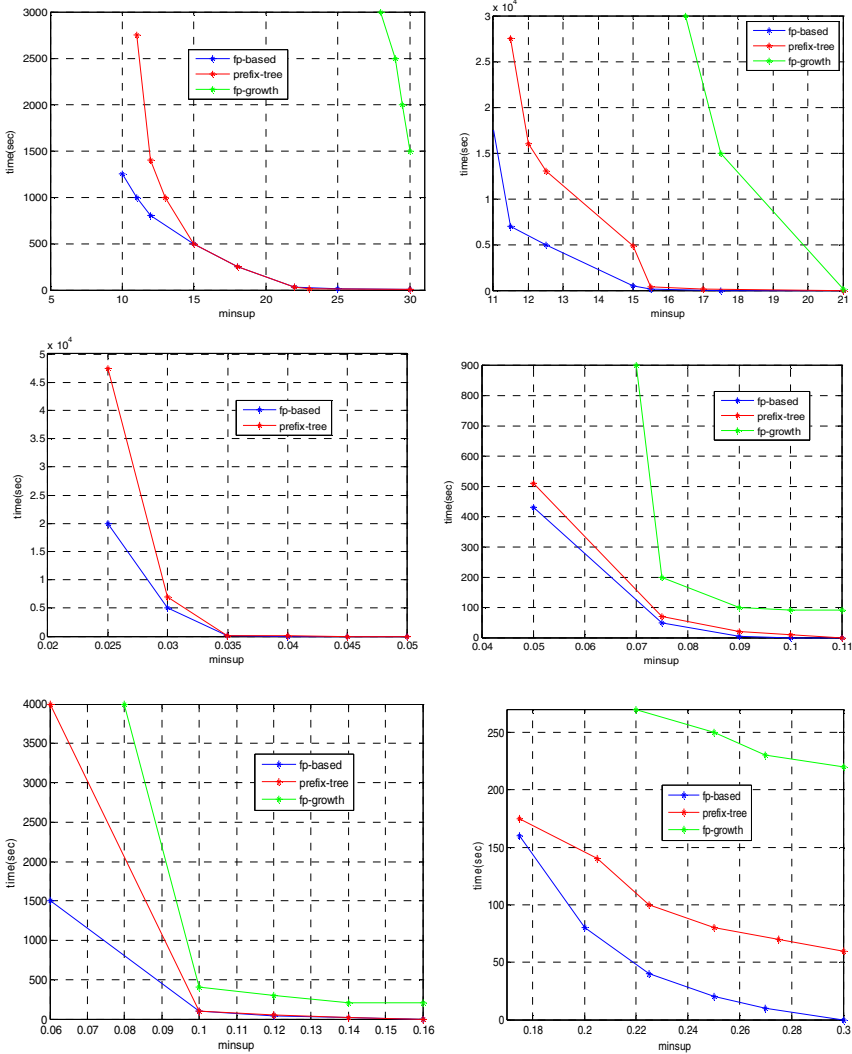


Fig. 3. Frequent WPs set extraction time for the FP-based algorithm. (a) Connect. (b) Pumsb. (c) Kosarak. (d) T10I200P20D2M. (e) T15I100P20C1D5M. (f) T20I100P15C1D7M.

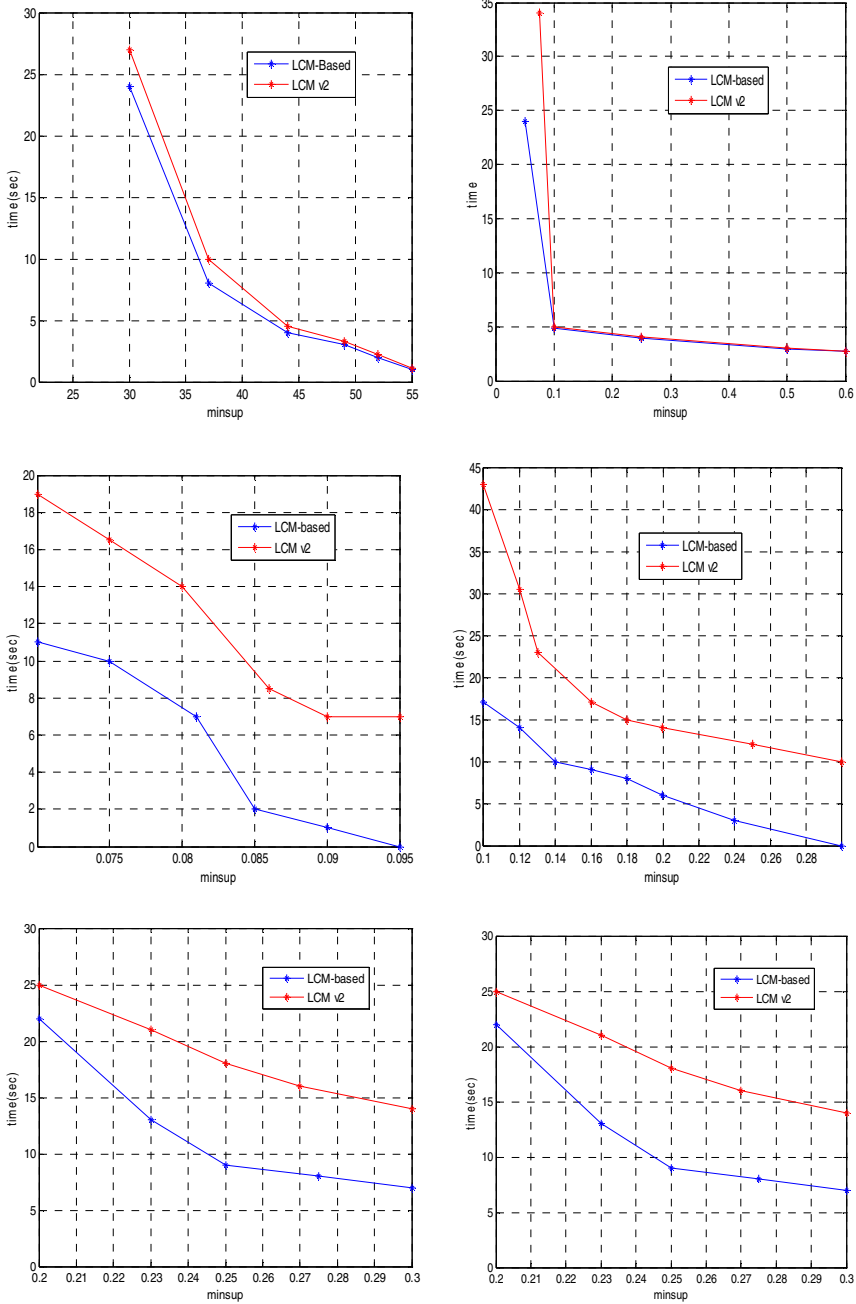


Fig. 4. Frequent web page set extraction time for the LCM-based algorithm (a) Connect. (b) Pumsb. (c) Kosarak. (d) T10I200P20D2M. (e) T15I100P20C1D5M. (f) T20I100P15C1D7M.

5 Conclusion

The WPs-Mine index provides a complete and compact representation of transactions in the database. Its novel index structure supports web page set mining into a relational database. It is a general structure that supports different algorithmic approaches to web page set extraction. The WPs-Hash index tree structure provides efficient access reducing I/O time. Performance is better than the state-of-the-art algorithm FP-growth, LCM v.2. Further extension may be to have still efficient compact structure for different data distributions and incremental updating of index. Incremental update is feasible without re-accessing the original transactional database. Incremental update can be done considering the transaction where session was ended.

References

1. Aggarwal, C., Yu, P.: A new framework for itemset generation. In: Proceedings of the PODS Conference, Seattle, WA, USA, pp. 18–24 (June 1998)
2. Agrawal, R., Imielinski, T., Swami, A.: Mining association rules between sets of items in large databases. In: Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, Washington, D.C, pp. 207–216 (May 1993)
3. Silverstein, C., Brin, S., Motwani, R.: Beyond market baskets: Generalizing association rules to dependence rules. *Data Mining and Knowledge Discovery*, 39–68 (1998)
4. Kotásek, P., Zendulka, J.: Comparison of Three Mining Algorithms for Association Rules. In: Proc. of 34th Spring Int. Conf. on Modeling and Simulation of Systems (MOSIS 2000), Workshop Proceedings Information Systems Modeling (ISM 2000), pp. 85–90 (2000)
5. Han, J., Pei, J., Yin, Y.: Mining frequent patterns Candidate generation. In: Proc. 2000 ACM-SIGMOD Int. Management of Data (SIGMOD 2000), Dallas, TX (2000)
6. Agrawal, R., Srikant, R.: Fast Algorithm for Mining Association Rules. In: Proc. 20th Int'l Conf. Very Large Data Bases, VLDB 1994 (September 1994)
7. Han, J., Pei, J., Yin, Y.: Mining Frequent Patterns without Candidate Generation. In: Proc. ACM SIGMOD (2000)
8. Uno, T., Kiyomi, M., Arimura, H.: LCM ver. 2: Efficient Mining Algorithms for Frequent/Closed/Maximal Itemsets. In: Proc. IEEE ICDM Workshop Frequent Itemset Mining Implementations, FIMI (2004)

Iris Liveness Detection for Semi-transparent Contact Lens Spoofing

Niladri B. Puhan¹, Sudha Natarajan², and A. Suhas Hegde¹

¹International Institute of Information Technology, Bangalore, India
nbpuhan@iiitb.ac.in, suhas.hegde@iiitb.net

²School of Computer Engineering, Nanyang Technological University, Singapore
sudha@ntu.edu.sg

Abstract. Park's proposed liveness detection technique in [5] detects the fake irises by checking the pupil size variation and textural feature change in local iris area. A new semi-transparent contact lens based spoofing method is designed to overcome the liveness technique. The spoofing method generates artificial texture image by replicating imposter's natural iris texture adjacent to pupil boundary. Simulation results show that by printing the artificial texture image on semi-transparent contact lens, an imposter can pass through the liveness detection to illegally make an entry despite being in negative-watch lists. A countermeasure algorithm is suggested for improving the liveness detection technique.

Keywords: iris recognition, liveness detection, spoofing, semi-transparent contact lens, security.

1 Introduction

Accumulated empirical evidence suggests the usefulness of human iris texture as a unique, stable and non-invasive biometric [1, 2]. Liveness detection, the act of determining vitality, is an important and challenging issue which determines the trustworthiness of biometric system security against spoofing. The spoofing methods include iris texture printed on paper and photographic surface, re-played video, fake glass/plastic eye and iris texture printed on contact lens (most difficult). The countermeasures include frequency spectrum analysis, red-eye effect, Purkinje reflections, hippus, eyelid movements and pupillary light reflex.

He et al. applied a combination of statistical features and SVM for detection of fake iris texture on opaque contact lens [3]. A new iris spoof detection method was proposed using boosted local binary patterns to discriminate between opaque contact lens and live iris textural patterns [4].

Checking the pupil size variation to surrounding illumination, the pupillary light reflex, is an effective way to detect spoofing. Fake irises such as the printed iris, photographs, plastic eye and opaque contact lens can be detected using this technique. Such fake iris images do not display the change in pupil's size. However, if the contact lens is partly opaque containing artificial texture (outer) and partly transparent

(adjacent to pupil), spoofing becomes hard to detect using pupillary light reflex. In this semi-transparent contact lens case, pupillary light reflex is visible through the transparent portion.

To overcome such spoofing, Park proposed an enhanced countermeasure of checking texture features in the local iris area adjacent of pupil boundary [5]. The iris pattern of live iris is dilated and contracted in case of pupil's movement. However, the fake iris is not dilated and contracted like that of live iris. So checking the iris feature changing in the local iris area is useful. Park divided the iris region (with light on) into 8 tracks and extracted 4 iris tracks adjacent to pupil boundary which are then converted into a rectangular image. Then, wavelet filtering is applied in the rectangular image to extract iris features. This operation is performed to extract iris features when light is off. If the iris feature values between two images are similar, the input iris image is regarded as live and vice versa.

In this paper, we define our scope in an iris recognition system which applies Park's liveness detection technique. An imposter who wants not to be recognized wears a semi-transparent contact lens containing artificial texture. The scenario is possible in security applications where the iris recognition system is used to prevent entry of unwanted persons by searching against negative watch-lists at ports of entry [2]. We propose a new spoofing method to design such artificial iris texture that can be printed on a semi-transparent contact lens. If the imposter wears the contact lens, it would be detected as live. Further, the imposter's correct identity would not be recognized due to different iris texture printed on the contact lens.

The paper is organized as follows. Section 2 presents the new semi-transparent contact lens based spoofing method. In Section 3, the performance of the proposed method is demonstrated by generating the artificial iris texture that is subjected to the liveness detection. Finally, some conclusions are given in Section 4.

2 Proposed Method

A schematic diagram of Park's liveness detection is presented in Fig. 1. The imposter wears a fake contact lens that is semi-transparent [Fig. 1(a)]. A part of the natural texture of imposter eye adjacent to pupil is visible initially when the light is on. In Fig. 1(a), the iris region has two types of textures: artificial texture and natural texture (shaded). When the light is off, the pupillary light reflex occurs and the pupil size changes by a certain amount.

We assume that only artificial texture is present after the pupillary light reflex [Fig. 1(b)]. This happens if the radius of the natural texture region is less than the pupil size variation. The circular iris regions in Fig. 1(a, c) are converted into the rectangular images by applying normalization, shown in Fig. 1(b, d). If top-halves of the rectangular images are texturally similar, the imposter could pass in the liveness detection.

Let the size of the rectangular image (I_1) be $M \times N$ pixels and the radius of the circular iris region is R in Fig. 1(a). Let the pupil of radius R_p expands by a factor of radius E pixels. A row of pixels in (I_1) corresponds to a ring of N pixels in the circular iris region. Thus, the rectangular image corresponds to a total of M rings in

the circular iris region. Similarly, the circular iris region of radius $R - E$ pixels is converted to the rectangular image (I_2) [Fig. 1 (c, d)]. In Fig. 1(b), let the no. of rows in (I_1) generated from the natural iris texture region be denoted as M_1 , shown as the shaded region (T_b) . The first row of (T_b) corresponds to a ring of N pixels surrounding the immediate periphery of the pupil in Fig. 1(a). Let the first M_1 rows in (I_2) be denoted as (T_a) .

For liveness detection, (T_a) and (T_b) should be texturally similar. It is then observed that the M_1 rows of (T_a) should be similar with $(M_1 + 1)$ to $(2M_1)$ row numbers in (I_1) . This comes from the geometrical relationship between the circular iris regions, where the artificial texture regions corresponding to both sets of row numbers in (I_1) and (I_2) are situated close to each other. Due to the liveness detection condition, the row numbers $(M_1 + 1)$ to $(2M_1)$ in (I_2) should be similar to the corresponding row numbers in (I_1) . The similarity observation due to the liveness detection and geometrical condition indicates that the required rectangular images (I_1) and (I_2) can be created by suitably replicating (T_b) . If the sub-images consisting of first $M/2$ rows of pixels of (I_1) and (I_2) becomes texturally similar, the input iris is decided to be live. In following, we describe the algorithm to generate the artificial texture region shown in Fig. 1(b).

Spoofing algorithm

Input: The imposter's original iris image

Output: The artificial texture image

1. Segment the circular iris region of radius R pixels
2. Normalize the segmented iris image to generate the rectangular image (I_1) of size $M \times N$ pixels
3. Decide the pupil expansion factor (E)
4. Compute M_1 and (T_b) from (I_1) using normalization model and the value of E
5. Compute (I_2) by replicating (T_b) using Eq. 1 and 2.
6. Generate the circular artificial iris texture region of radius $(R - E)$ pixels from (I_2) by applying inverse normalization procedure.

A countermeasure to the proposed spoofing method is to find, whether there exists a replication of texture in the artificial texture image. Since the iris texture is random by nature, such replication will not be found in live iris images. In following, an algorithm is proposed to be incorporated in Park's liveness detection to detect the proposed spoofing method.

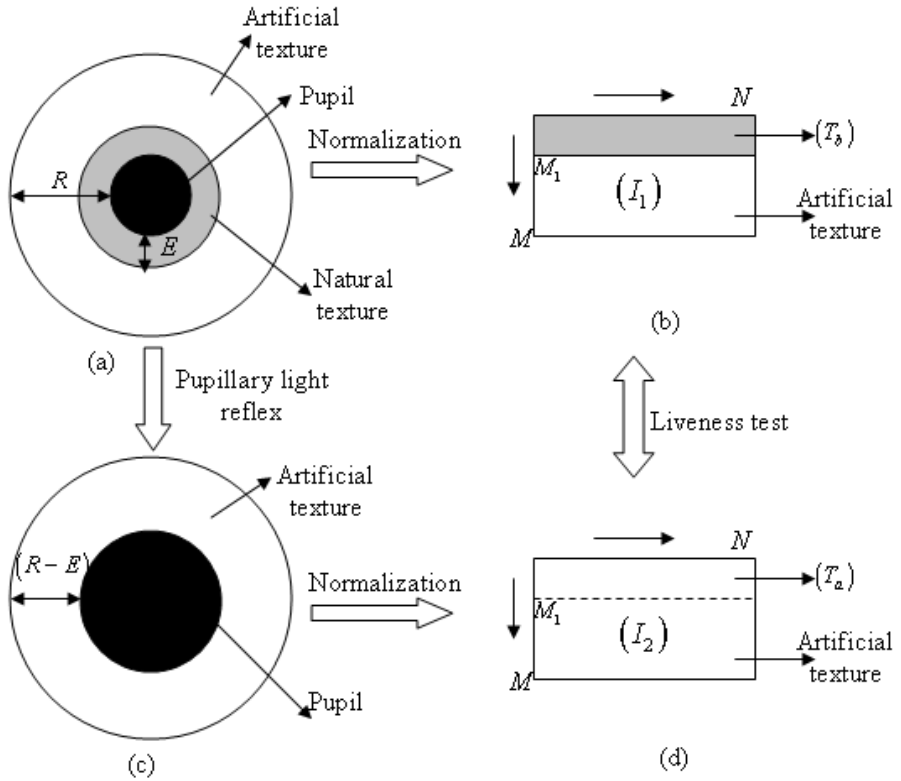


Fig. 1. Schematic diagram of Park's liveness detection technique using pupillary light reflex

Countermeasure algorithm

Input: The iris images before and after pupillary light reflex (I_a, I_b)

Output: Decide whether the input iris image is live

1. Segment and normalize the iris images (I_a, I_b) to (I_a^n, I_b^n)
2. Using the normalization model and the pupil expansion radius, compute M_1
3. Compute the textural correlation between two sub-images, derived from 1 to M_1 and $(M_1 + 1)$ to $2M_1$ rows of pixels in I_b^n
4. If the correlation value is found to be high, decide the input iris image as fake

3 Results and Discussion

We demonstrate the effectiveness of the proposed method by generating artificial texture images. The original iris images are obtained from the UPOL iris database [6]. The original segmented image is shown in Fig. 2(a). When the image belongs to an

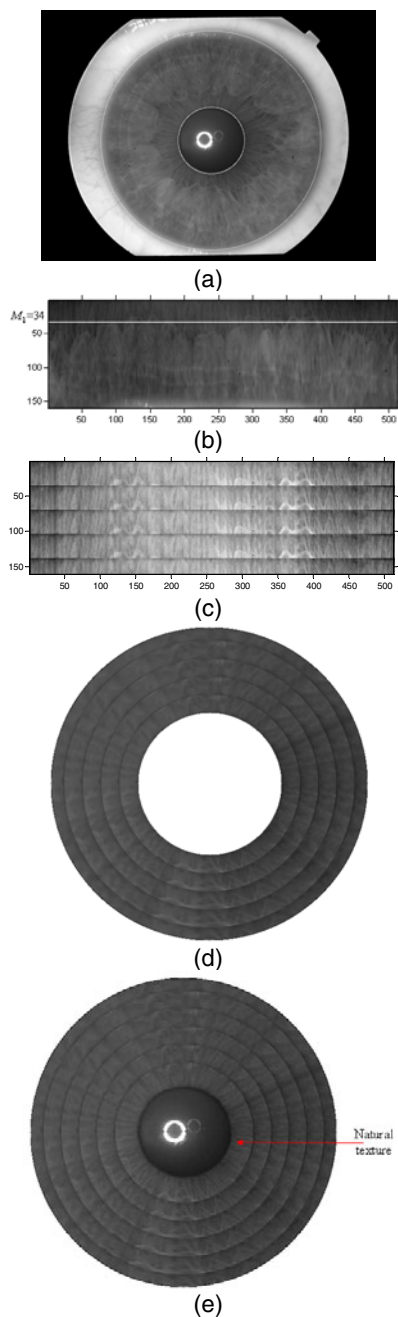


Fig. 2. (a) Segmented iris image, (b) normalized image, (c) normalized image after the replication of natural texture region, (d) artificial texture image, (e) simulated fake iris image before pupillary light reflex, (f) simulated fake iris image after pupillary light reflex, (g) top-half of the normalized image of (e) and (h) top-half of the normalized image of (f).

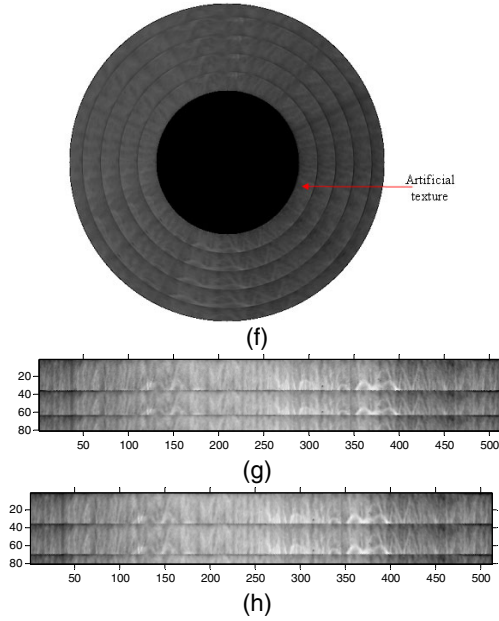


Fig. 2. (Continued)

imposter, the goal is to generate an artificial iris texture image to pass the liveness detection. The iris radius (R) and pupil radius (R_p) is found to be 176 and 76 pixels respectively. In Fig. 2(b), the segmented iris image is normalized to a rectangular image (I_1) of size 160×512 pixels using the rubber-sheet model [1, 7]. The pupil expansion factor (E) of 38 pixels is chosen which is 50% of the given pupil radius. In Fig. 2(b), the texture region (T_b) is indicated in (I_1) where the value of M_1 is 34. In Fig. 2(c), the rectangular image (I_2) is generated by replicating the texture region (T_b). The replicated image (I_2) consists of k blocks of (T_b) and a last block with first l rows of (T_b). The values of k and l are given according to Eq. 1 and 2.

$$k = \left\lfloor \frac{M}{M_1} \right\rfloor \tag{1}$$

where $\lfloor \cdot \rfloor$ is the floor operation.

$$l = M - kM_1 \tag{2}$$

The circular artificial texture image is generated from (I_2) by applying the rubber-sheet model. The pixel values at a total of 81,920 (MN) coordinates of 160 (M)

circular rings in the artificial texture region are obtained from (I_2). The cubic surface interpolation technique is used to estimate the integer coordinate pixel values [8]. We apply the MATLAB command ‘*griddata*’ to implement the interpolation technique. In Fig. 2(d), the artificial iris texture image is shown that can be printed on a semi-transparent contact lens. When the imposter wears the contact lens, the iris will be presented to the acquisition system as shown in Fig. 2(e). After the pupillary light reflex, the iris image will be modified to Fig. 2(f).

We apply the liveness detection between the images in Fig. 2(e, f) which are the simulated fake iris images before and after pupillary light reflex. The normalized sub-images of size 80×512 pixels in Fig. 2(g, h) are texturally compared. The correlation between the sub-images is quantified in terms of normalized Hamming distance (HD_{live}) between the iris code templates [7]. In the feature extraction technique, the phase data from 1D Log-Gabor filters is extracted and quantized to four levels to encode a binary iris code template. The value of HD_{live} is found to be 0.17, which shows good textural similarity. The result is expected since the artificial texture image was particularly generated to pass the liveness detection.

To verify if the imposter can be recognized, we compare imposter’s original iris image (which is in the negative watch-list) and the iris image captured after wearing the contact lens [Fig. 2(a, e)]. After segmentation and normalization to the rectangular images of size 160×512 pixels, the feature extraction technique [7] is used to compute the normalized Hamming distance (HD_{recog}). The value of HD_{recog} is found to be 0.38. In [2], the smallest and largest Hamming distances found in the set of 200 billion comparisons were around 0.26 and 0.75 respectively. If the threshold between intra-class and inter-class distribution is chosen as 0.26, the fake iris would be decided not to be from the imposter. Thus, the imposter would be declared not to be in the negative-watch list and cleared for entry.

Table 1. The values of HD_{live} , HD_{recog} and HD_{CM} for 10 original iris images

Image no	HD_{live}	HD_{recog}	HD_{CM}
1.	0.19	0.41	0.03
2.	0.21	0.41	0.03
3.	0.17	0.40	0.02
4.	0.17	0.42	0.02
5.	0.20	0.41	0.03
6.	0.24	0.44	0.04
7.	0.20	0.41	0.03
8.	0.17	0.39	0.02
9.	0.24	0.43	0.03
10.	0.21	0.40	0.03

We generate artificial texture images for 10 iris images using the proposed spoofing method. In Table 1, the values of HD_{live} and HD_{recog} are given for all cases. The low values of HD_{live} show that the fake iris images would be detected as live. The high values of HD_{recog} show that the recognition system would fail to identify the imposter. In order to validate the proposed countermeasure, we implement the countermeasure algorithm for the fake iris images such as in Fig. 2(e, f). We compute the textural correlation (in step-3) using the normalized Hamming distance (HD_{CM}), similar to HD_{live} and HD_{recog} [7]. The low values of HD_{CM} in Table 1 show that the input iris images contain a replication of texture; thus they are the fake irises.

4 Conclusion

In this paper, we proposed a new semi-transparent contact lens based spoofing method that can overcome existing liveness detection technique. The artificial texture is generated by replicating the natural texture region adjacent to pupil boundary. After wearing the contact lens, the iris image of the imposter would be detected as live. Due to printed iris texture, the imposter would be able to hide own identity from negative watch-lists. A countermeasure is suggested to improve Park's liveness detection technique against the spoofing method.

References

1. Daugman, J.: High Confidence Visual Recognition of Persons by a Test of Statistical Independence. *IEEE Trans. Pattern Analysis and Machine Intelligence* 15(11), 1148–1161 (1993)
2. Daugman, J.: Probing the uniqueness and randomness of IrisCodes: Results from 200 billion iris pair comparisons. *Proc. of the IEEE* 94(11), 1927–1935 (2006)
3. He, X., Shi, A.S., Lee, P.C.: Statistical texture analysis based approach for fake iris detection using SVM. In: *IEEE Int. Conf. on Biometrics*, pp. 540–546 (2007)
4. He, Z., Sun, Z., Tan, T., Wei, Z.: Efficient iris spoof detection via boosted local binary patterns. In: Tistarelli, M., Nixon, M.S. (eds.) *ICB 2009*. LNCS, vol. 5558, pp. 1080–1090. Springer, Heidelberg (2009)
5. Park, K.R.: Robust fake iris detection. In: Perales, F.J., Fisher, R.B. (eds.) *AMDO 2006*. LNCS, vol. 4069, pp. 10–18. Springer, Heidelberg (2006)
6. Dobeš, M., Machala, L.: Iris Database, <http://www.inf.upol.cz/iris/>
7. Masek, L., Kovesi, P.: *MATLAB Source Code for a Biometric Identification System Based on Iris Patterns*. University of Western Australia (2003)
8. Watson, D.F.: *Watson: Contouring: A guide to the analysis and display of spatial data*. Pergamon, Oxford (1994)

Stego-Image Generator (SIG) - Building Steganography Image Database

P. Thiagarajan, G. Aghila, and V. Prasanna Venkatesan

CDBR-SSE Lab, Department of Computer Science,
Pondicherry University, Puducherry 605 014
{thiyagu.phd, aghilaa}@gmail.com,
prasanna_v@yahoo.com

Abstract. Any Universal Steganalysis algorithm developed should be tested with various stego-images to prove its efficiency. This work is aimed to develop a tool to build the stego-image database which is obtained by implementing various RGB based Least Significant Bit Steganographic algorithms. Though there are many stego-images sources available on the internet it lacks in the information such as how many rows has been infected by the steganography algorithms, how many bits have been modified and which channel has been affected. These parameters are important for Steganalysis algorithms to rate its efficiency. For the experiments conducted using proposed method, Stego-Image Image Generator Tool, images are chosen from board categories such as animals, nature, person etc to produce variety of Stego-Image. The proposed Stego-Image Generator (SIG) has been compared with various Stego-tools against various parameters to prove its efficiency.

Keywords: Image Steganography, Steganalysis, Forensics Examiner, LSB, RGB, Stego-Images.

1 Introduction

Steganography is the practice of concealing the very presence of message during communication [1]. Like two sides of the coin, Steganography has both advantages and disadvantages. It depends on the person who uses it for example it is in the hands of the scientist he may use it for the military purpose or if it is in the hands of the terrorists he may use image steganography to snatch the attack plan among his team members secretly and communicate via internet. In the later case it is more important for cyber crime investigators to detect the stego-image which is used by terrorists. The technique which is used to identify the images that contains the secret message is called as Steganalysis. Our work aims to build variety of stego-images which is useful for steganalyst to test their Steganalysis algorithms. Section 2 gives outline about Steganography, Section 3 gives brief outline about Steganalysis, Section 4 outlines about the various Image Steganography tools and the need for proposed SIG tool, Section 5 describes the Architecture of proposed SIG tool, Section 6 compares the

SIG tool with various other Stego-tool reported in the literature Survey and Section 7 concludes the paper.

2 Steganography

Steganography can be best explained through prisoner’s problem Alice wishes to send a secret message to Bob by hiding information in a clean image. The stego image (clean image + secret message) passes through Wendy (a warden) who inspects it to determine if there is anything suspicious about it. Wendy could perform one or several scan to decide if the message from Alice to Bob contains any secret information. If the decision is negative then Wendy forwards the message to Bob—Wendy acts as a passive warden. On the other hand, Wendy can take a conservative approach and modify all the messages from Alice to Bob irrespective of whether any information is hidden by Alice or not [3].

In this case, Wendy is called an active warden. Wendy will have constraints such as the maximum allowable distortion when modifying the message etc. For example, if the clean messages are digital images, then Wendy cannot modify the stego message to an extent that perceptually significant distortions are induced. Fig 1 represents the pictorial representation of how message is embedded and extracted in steganography.

LSB methods are most commonly used steganography techniques to embed in cover image. Least significant bit of some or all of the bytes inside an image is changed to a bit of the secret message. When using a 24-bit RGB Color image, a bit of each of the red, green and blue color components can be used. In other words, one pixel can store 3 bits of secret message.

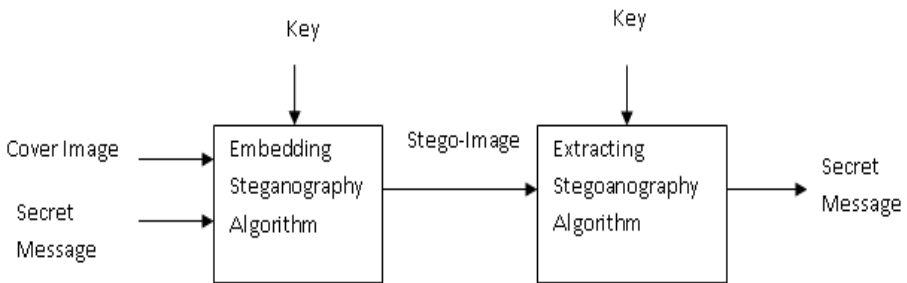


Fig. 1. Steganography Process Diagram

Components

- Secret Message - The message to be embedded
- Cover Image – An image in which Secret Message will be embedded.
- Stego Image - Cover image that contain embedded message.
- Key – Additional data that is needed for embedding and extracting process.

- Embedding Steganography Algorithm - Steganography Algorithm used to embed secret message with cover image.
- Extracting Steganography Algorithm - Inverse function of embedding, in which it is used to extract the embedded message (secret message) from stego image.

3 Steganalysis

Steganography can be applied in numerous fields such as authentication, secret communication in military, banking etc., however it also depends on the person who is using. There are strong indications that steganography has been used for planning criminal activities [2]. In this way, it is important to detect the existence of hidden messages in digital files. As with cryptography and cryptanalysis, Steganalysis is defined as the art and science of breaking the security of steganography systems. The goal of steganography is to conceal the existence of a secret message. While the goal of Steganalysis is to detect that a certain file contains embedded data. The stegosystem can be extended to include scenarios for different attacks [4] similar to the attacks on cryptographic systems. In general, extraction of the secret message could be a harder problem than mere detection. The challenges of steganalysis are listed below

- a) To get Stego-Image Database
- b) To test the Steganalysis algorithm against different payload Stego-Images to check its robustness
- c) To test the Steganalysis algorithm from various categories of images such as animals, fruits, natural scene etc.,
- d) Identification of embedding algorithm
- e) Detection of presence of hidden message in cover image
- f) Estimation of embedded message length
- g) Prediction of location of hidden message bits
- h) Estimation of secret key used in the embedding algorithm
- i) Extraction of hidden message

In this paper we address the first three issues of Steganalysis by generating different stego-images using the proposed SIG (Stego-Image Generator) tool.

4 Stego-Image Tools Reported in Literature Survey

There are many steganography tools available [5] from that we have surveyed 15 Stego-image generation tools. Detailed survey of the below tools have been made and they have been classified according to the format of stego-image it produces, availability of code and released or recently updated year. Table 1 gives details about the various stego-image generation tools where the code of the tool is available [6] [7] [8] [9] [10] [11]. Table 2 gives the detail about the Stego-image generation tool where the code is not available but its executable is available [12] [13] [14] [15] [16] [17].

Table 1. Image Steganography Tools with Source Code (NA- Not Available)

S.No	Name of Image Steganography Tool	Released Year / Updated Year	Format	Availability of Code
1	Blind slide	NA	BMP	Yes
2	Camera Shy	2002	JPEG	Yes
3	Hide4PGP	2000	BMP	Yes
4	JP Hide and Seek	1999	JPEG	Yes
5	Jsteg Jpeg	1995	JPEG	Yes
6	Mandelsteg	1995	GIF	Yes
7	Steghide	2003	BMP	Yes
8	wbStego	NA	BMP	Yes

Table 2. Image Steganography Tools without Source Code (NA- Not Available)

S.No	Name of Image Steganography Tool	Released Year / Updated Year	Format	Availability of Code
1	Camouflage	NA	PNG	No
2	Hide & Seek	1995	BMP,GIF	No
3	S-Tools	2000	BMP	No
4	Steganos	1996	BMP	No
5	StegMark	2007	BMP,GIF,PNG	No
6	Invisible Secrets	NA	BMP,JPEG,GIF	No
7	Info Stego	2007	BMP,JPEG,PNG	No

All the above tools mentioned in table 1 and 2 does not important input parameters in the tools such as how many LSB (Least significant Bits) to be replaced, channel in which the embedding should be done and number of rows should to be affected This section clearly depicts the need for such tool which gets the above desired input from the user and with the help of these inputs user sort the stego-images generated from the proposed tool.

5 Architecture of SIG Tool

Though there are many tools available for stego-image generation certain important features are missing

- Lack of getting user desired inputs such as channels in which the changes to be incorporated, number of bits to be affected, number of rows to be affected.
- Proper Classification of Stego-images with respect to number of rows affected

- c) Proper Classification of Stego-images with respect to number of bits changed
- d) Proper Classification of Stego-images with respect to number of channels changed

To overcome the above issues a new architecture was proposed for Stego-Image Generator (SIG) tool.

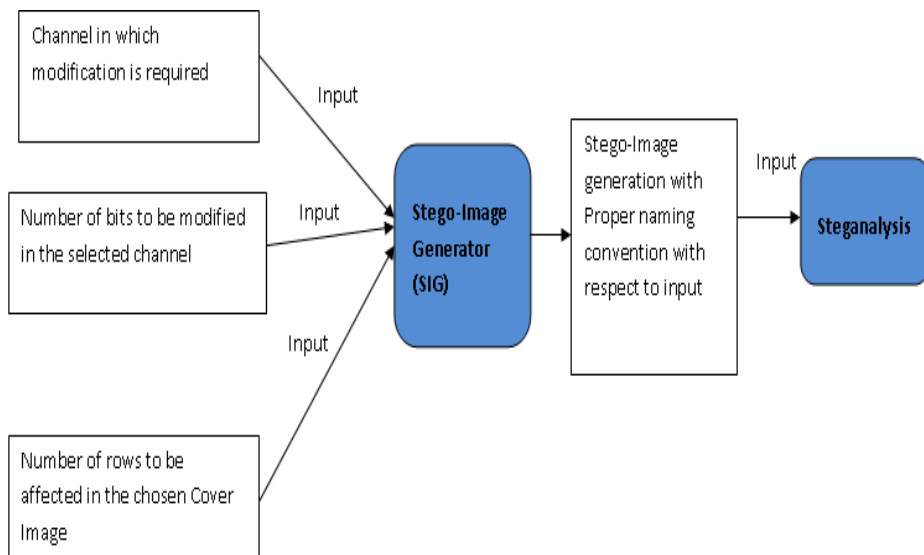


Fig. 2. Architecture of Stego-Image Generator

The above architecture in figure 2 was designed specifically for the RGB based colour Images. The SIG Tool accepts the input such as number of bits to be replaced, channel in which the modification is done and number of rows to be affected. These inputs are taken by the SIG tool and produce the stego-image with proper naming convention so that it can be sorted accordingly. This tool can generate 63 different kinds of Stego-Images from a single Image. The calculation is shown below

In 7 different ways RGB channels can be chosen such as R, G, B, RG, RB, GB and RGB. In 3 different ways the Rows can be selected such as 5, 10, 20 rows. In 3 different ways the Least Significant Bits are changed such as 1, 3 or 4 bits. Reason for choosing maximum 20 rows is that in steganography message communicated will not be of much length which can be accommodated with in 20 rows. In Stego-Image maximum of 3 to 4 bits are changed in the LSB of every pixel. 63 Stego-Image can be obtained from a single Cover-Image and the calculation is shown below

Number of ways in which RGB Channels can be chosen = 7

Number of ways in which the rows can be chosen = 3

Number of ways in which the number of bits that need to be changed to be chosen = 3

Total number of Stego-Images generated for single Cover Image = $7*3*3 = 63$

Algorithm Used in Stego-Image Generator

Input: Cover Image, Input_Channel, No of Rows to be affected (Input_Rows), No of bits to be change (Input_Bits)

Output: Stego-Image

Algorithm:

For 1 to Input_Row

 For 1 to last_col

 Get the Input_channel and change the

 LSB_Bits according to the Input_Bits

 End

End

Save the generated Stego_Image with the meaningful naming convention such as

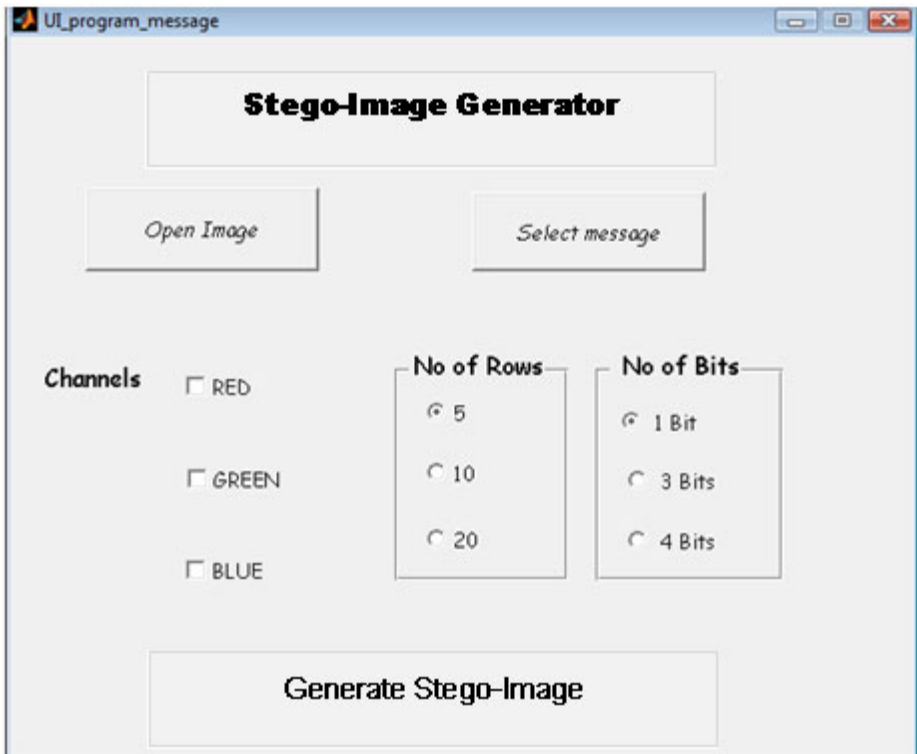
“filename_Inputbits_InputRows_InputChannel”

The above LSB algorithm used in Stego-Image Generator covers most of the RGB Steganography algorithm such as Pixel Indicator High capacity Technique Algorithm, RGB Intensity Image stego Algorithm etc. These 63 RGB Stego-Images generated by SIG tool forms the superset of RGB Stego-images. Therefore any RGB Steganalysis algorithm developed can extensively test with database obtained by SIG tool to test their robustness of the algorithm. So far 100 Cover-Images from different categories have been given as input to SIG tool. Images from different categories have been chosen to have various color combination in the Cover-Image. Some of the sample input from different category is shown in table 3. Currently SIG database contains 6300 Stego-Images.

The proposed Stego-Image Generator was implemented in Matlab 7 and the sample screen shots are shown in Figure 3 and 4. Figure 5 and 6 depicts the sample cover image and the stego-image obtained from Stego-Image Generator tool. The 63 Stego-images obtained from the single cover image using SIG tool is shown in Figure 7.

Table 3. Input Image Category for SIG Tool

S.No	Image Name	Category
1	Lotus.bmp	Flora
2	Monkey	Fauna
3	Baby	People
4	Sea	Natural
5	Cupcakes	Eatables
6	Tajmahal	Building

**Fig. 3.** GUI Stego-Image Generator

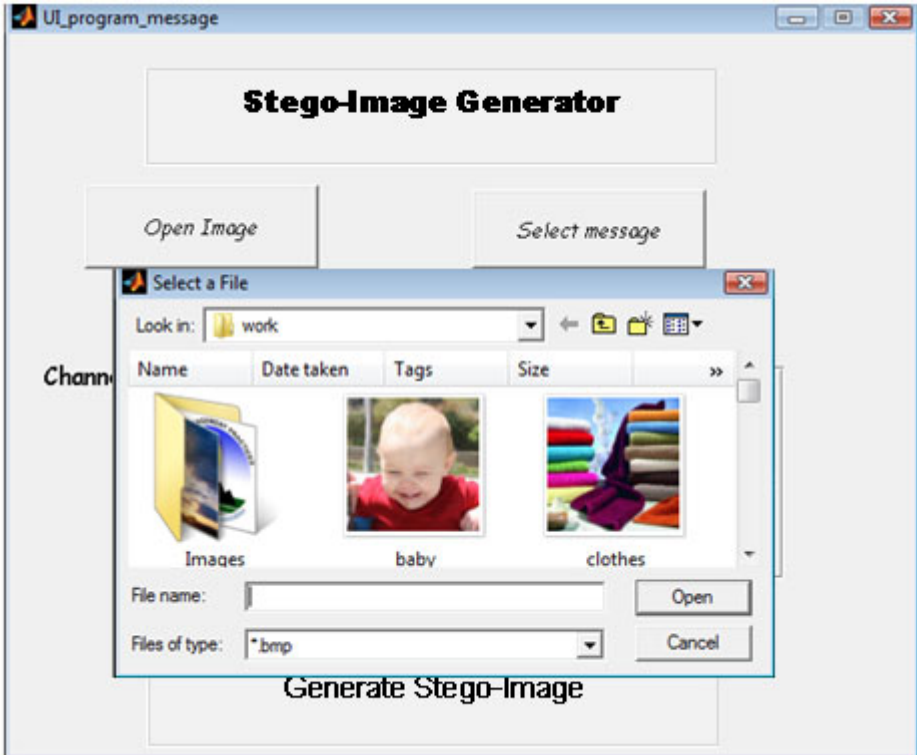


Fig. 4. Selecting Cover Image for Stego-Image Generator



Fig. 5. Cover Image



Fig. 6. Stego Image generated by replacing 3 bits from Red Channel in 5 rows

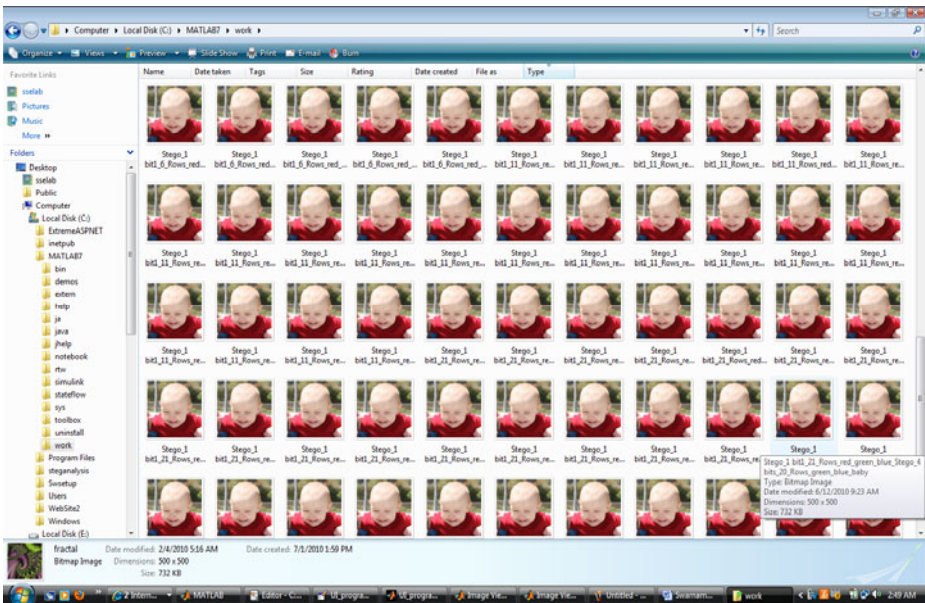


Fig. 7. Sixty three (63) Stego-Images obtained from baby image using SIG Tool

6 Comparison of Stego-Image Generator with other Tools

There are numerous tools available either in free or licensed version in the Internet for the production of Stego-Images. Mother of all Steganography algorithms in spatial domain is Least Significant Bit (LSB) algorithm. Our proposed SIG tool uses the LSB algorithm in RGB based colour images combined with choice in selecting the no of

LSB bits to be replaced, choice in selecting the colour channel, choice in selecting the no of rows to be affected. In table 4 the proposed SIG tool is compared with various Stego-Tools against various parameters.

S.No	Tools	Input such as Channel, Number of LSB to be replaced in Cover Image	Format	Number of Stego-Images Produced on per Cover Image
1	S-Tools	No	BMP	1
2	Camera Shy	No	JPEG	1
3	Steganos	No	BMP	1
4	Steghide	No	BMP	1
5	Info Stego	No	BMP,JPEG,PNG	1
6	SIG (Stego-Image Generator)	Yes	All format expect compressed image format	63

Stego-Image Generator (SIG) produces stego-images of all types expect compressed images type. All the Stego-Tool reported in the literature survey produces 1 stego-Image for the given cover Image. SIG Tool produces 63 different types of Stego-Images.

7 Conclusion

In this paper 15 different steganography tools have been surveyed and the areas in which they lack have been identified. To overcome the limitations in the existing system the SIG architecture has been proposed and implemented. This tool is very useful for forensic investigators and security agencies across the globe that is constantly analysing the Internet to detect communication that involves Image steganography. In the future we would test all these tools to develop robust real time Image steganalysis algorithm that can filter Steganography images which is used by terrorists for illegal activities.

Acknowledgement. This research, CDBR Smart and Secure Environment, was sponsored by National Technical Research Organisation (NTRO) and their support is greatly acknowledged. I would also like to thank Ms.G.Swarnambigai for her support in implementation.

References

- Berghel, H., O’Gorman, L.: Protecting ownership rights through digital watermarks. IEEE Computer 29(7), 101–103 (1996)
- Westfeld, A., Pfitzmann, A.: Attacks on Steganographic Systems. In: Proceedings of the Third International Workshop on Information Hiding, September 29-October 01, pp. 61–76 (1999)

3. Artz, D.: Digital Steganography: Hiding Data within Data. IEEE Internet Computing: Spotlight, 75–80 (May-June 2001)
4. Johnson, N., Jajodia, S.: Exploring steganography: Seeing the unseen. IEEE Computer 31(2), 26–34 (1998)
5. Pedram Hayati, A., et al.: Survey of Steganographic and Steganalytic Tools for the Digital Forensic Investigator. In: Workshop of Information Hiding and Digital Watermarking to be held in Conjunction with IFIPTM, Moncton, New Brunswick, Canada (July 2007)
6. Camera Shy, <http://hacktivism.com/projects/index.php>
7. Hide4PGP, <http://www.heinz-repp.onlinehome.de/Hide4PGP.htm>
8. Hide, J.P., Seek.: <http://linux01.gwdg.de/~alatham/stego.html>
9. Jsteg Jpeg, <http://www.nic.funet.fi/pub/crypt/steganography/>
10. Mandelsteg, <ftp://ftp.funet.fi/pub/crypt/steganography/>
11. Steghide, <http://steghide.sourceforge.net/>
12. Camouflage, <http://camouflage.unfiction.com/>
13. Hide, Seek:
<ftp://ftp.funet.fi/pub/crypt/steganography/hdsk41.zip>
14. S-Tools, <ftp://ftp.funet.fi/pub/crypt/mirrors/idea.sec.dsi.unimi.it/code/>
15. Steganos, <http://www.steganography.com/>
16. StegMark, <http://www.datamark-tech.com/index.htm>
17. Info Stego, <http://www.antiy.net/infostego/>

Review on OCR for Handwritten Indian Scripts Character Recognition

Munish Kumar¹, M.K. Jindal², and R.K. Sharma³

¹Assistant Professor, Computer Science Department,
GGS College for Women, Chandigarh, India

²Associate Professor, Department of Computer Science & Applications,
Panjab University Regional Centre, Muksar, India

³Professor, School of Mathematics & Computer Applications,
Thapar University, Patiala, India

munishcse@gmail.com, manishphd@rediffmail.com,
rksharma@thapar.edu

Abstract. Natural language processing and pattern recognition have been successfully applied to Optical Character Recognition (OCR). Character recognition is an important area in pattern recognition. Character recognition can be printed or handwritten. Handwritten character recognition can be offline or online. Many researchers have been done work on handwritten character recognition from the last few years. As compared to non-Indian scripts, the research on OCR of handwritten Indian scripts has not achieved that perfection. There are large numbers of systems available for handwritten character recognition for non-Indian scripts. But there is no complete OCR system is available for recognition of handwritten text in any Indian script, in general. Few attempts have been carried out on the recognition of Devanagari, Bangla, Tamil, Oriya and Gurmukhi handwritten scripts. In this paper, we presented a survey on OCR of these most popular Indian scripts.

Keywords: OCR, Handwritten character recognition, online, offline, Indian scripts.

1 Introduction

Nowadays, world is being influenced a lot by computers and almost all the important processing is being done electronically. As such, it becomes important that transfer of data between human beings and computers is simple and fast. Character recognition is a research problem that has been ongoing since the sixties. It stills an active area of research because the problem is complex in nature. Optical Character Recognition (OCR) is the most essential part of document analysis system. OCR is the field of pattern recognition, image and natural language processing. OCR is the recognition of printed or handwritten text by a computer. This recognition gives a significant benefit in order to bridge the gap between man and machine communication. The document analysis and recognition has played and currently playing a major role in pattern recognition research. In general, research on optical character recognition for Indian scripts is ongoing. But till now no solution has been offered that solves the problem

correctly and efficiently. The process of character recognition can be divided into two parts, namely, printed and handwritten character recognition. The printed documents can further be divided into two parts: good quality printed documents and degraded printed documents. Handwritten character recognition has been divided into offline and online character recognition, as shown in figure 1.

Offline documents are scanned images of prewritten text, generally on a sheet of paper. In online handwriting recognition, data are captured during the writing process with the help of a special pen and an electronic surface. Recognition of offline handwritten documents has been an active research area in the field of pattern recognition. Over the last few years, the numbers of laboratories all over the world are involved in research on handwriting recognition.

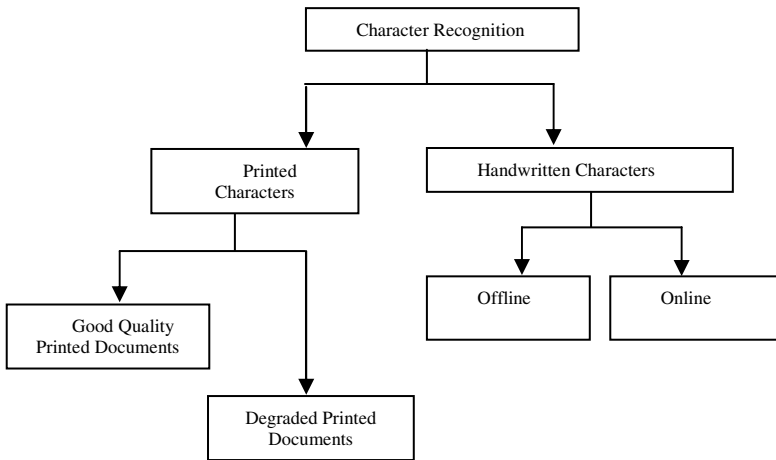


Fig. 1. Different character recognition systems

The recognition of cursive handwriting is very difficult due to large number of variations in shapes and overlapping of characters. In the offline handwriting recognition system, the pre-written document is converted into a digital image through optical scanner. In the handwritten script, there are variations in writing style, size *etc.* Handwritten Character Recognition (HCR) system will enable the computer to process the handwritten documents, which are currently processed manually. One can find these handwritten documents at various places such as post offices, banks, insurance offices and colleges *etc.* for processing data.

1.1 Stages of HCR

A complete process is followed for handwritten character recognition which is shown in figure 2.

1.1.1 Digitization

Digitization is the process whereby a document is scanned and an electronic representation of the original, in the form of a bitmap image, is produced. Digitization produces the digital image, which is fed to the pre-processing phase.

1.1.2 Preprocessing

Preprocessing is used for skew detection/correction, skeletonization, and noise reduction/removal. Skewness refers to the tilt in the bit mapped image of the scanned paper for OCR. It is usually caused if the paper is not fed straight into the scanner. Skeletonization is used for decreasing the line width of text from many pixels to single pixel. Noise removal is used to remove unwanted bit pattern which does not play any significant role in document.

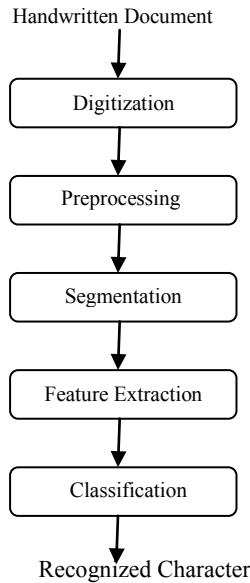


Fig. 2. Handwritten character recognition system.

1.1.3 Segmentation

In the character recognition, segmentation is very important for recognition. Segmentation is used to break the script into lines, words and characters. The challenge of a segmentation technique lies in the decision of best segmentation point for line, word and character isolation. In handwritten script, lots of features are available for segmentation provided by skeleton and the intersection point. Incorrect segmentation can lead to the incorrect recognition. Segmentation of handwritten text is a difficult task owing to variety of writing styles.

1.1.4 Feature Extraction and Classification

Feature extraction is the phase which is used to measure the relevant shape contained in the character. In the feature extraction phase, one can extract the features according to levels of text, e.g., character level, word level, line level and paragraph level. The classification phase is the decision making phase of an OCR engine, which uses the features extracted in the previous stage for making the class memberships in pattern recognition system. The preliminary aim of classification phase of OCR is to develop the constraint for reducing the misclassification relevant to feature extractions.

2 Properties of Indian Scripts

There are 23 languages in India [1] namely Assamese, Bengali, Bodo, Dogri, English, Gujarati, Hindi, Kannada, Kashmiri, Konkani, Maithili, Malayalam, Manipuri, Marathi, Nepali, Oriya, Punjabi, Sanskrit, Santhali, Sindhi, Tamil, Telugu and Urdu. There are 14 different scripts in India Assamese, Bangla, Devanagri, Gujarati, Gurmukhi, Kannada, Kashmiri, Malayalam, Oriya, Roman, Tamil, Telugu and Urdu used for writing these languages. Indian scripts are different from non-Indian scripts in several ways. Indian scripts are composition of symbols like: consonants and modifiers. In Indian scripts case sensitivity is absent. The Indian scripts are divided into three zones as shown in figure 3. As compared to non-Indian scripts, the research on OCR of handwritten Indian scripts has not achieved that perfection. Few attempts have been carried out on the recognition of Devanagari, Bangla, Tamil, Oriya, Telugu and Gurmukhi handwritten scripts [2-5]. There is no complete OCR system is available for commercial use for recognition of handwritten text in any Indian script, in general.

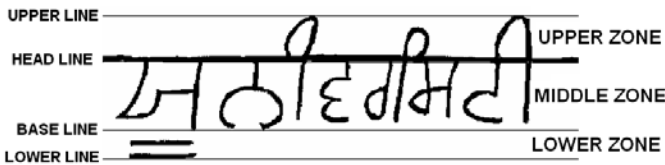


Fig. 3. Different zones of Gurmukhi text

3 Handwritten Character Recognition

The research for developing an OCR system started in the nineteenth century and a system was available in 1929. Modern version of OCR was developed in 1951 by David Shepard. This became very popular owing to its commercial use. Handwriting is the natural mode of collection and storing information for the human beings. It can also be used for communication between human beings and computers. Handwritten character recognition system can be evaluated from several perspectives. One major distinction can be made from the recognition process either during the writing (on-line) or from earlier handwritten document (offline). The recognition of handwritten character is very complex due to non-uniformity in size and style. In general, the location of the characters is not predictable, nor is the spacing between them. An online handwriting OCR was also available during 1950s in which an electronic tablet was used to capture the x-y values of pen movement. Mori *et al.* [6] have divided the OCR into three generations: first generation is for printed documents, second generation for handwritten documents and third generation is for degraded printed text and documents with text, graphics and mathematical symbols.

3.1 Recognition of Devanagari Script

Devanagari, an alphabetic script, is used by a number of Indian languages, including Sanskrit, Hindi and Marathi. Kumar [7] has proposed an AI based technique for machine recognition of Devanagari handwritten script. He has used three levels of

abstractions to describe this technique. Recognition of Devanagari handwritten script was provided in 2007 by Hanmandlu *et al.* [8]. They have proposed the system for recognition of handwritten Hindi characters based upon the membership function of fuzzy sets. Sethi and Chatterjee [9] have also done some work on Devanagari script. They have presented a Devanagari hand-printed numeral recognition system based on binary decision tree classifier. Bansal and Sinha [5] have proposed the technique for complete Devanagari script recognition. In this research, they recognize the character into two steps, in the first step, they recognize the unknown stroke and in second step the character based on these strokes is recognized. Bajaj and Chaudhury [10] have proposed a system for hand-written numeral recognition of Devnagari characters.

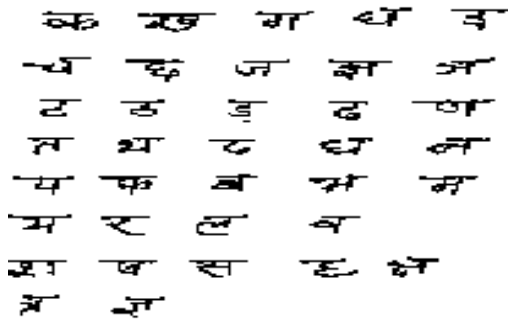


Fig. 4. Handwritten Devanagari characters

3.2 Recognition of Bangla Script

The maximum work for recognition of handwritten characters has been done on Bangla characters. In 1982, Chaudhury *et al.* [11] have proposed a recognition scheme using a syntactic method for connected Bangla handwritten numerals. In this system, the skeleton i.e. structure of character is matched. Pal *et al.* [12] have proposed the technique for Bangla handwritten pin code recognition system. They input the bitmap document and then water reservoir concept is applied to segment and recognize the pin code.



Fig. 5. Handwritten Bangla characters

Bishnu and Chaudhuri [13] have proposed technique for segmentation of Bangla handwritten text into characters based on recursive shape. Dutta and Chaudhuri [14] have developed an isolated optical character recognition system for Bangla alphabets and numerals using curvature features. Pal *et al.* [15] have proposed the technique for Bangla handwritten numerals document and then water reservoir concept is applied to segment. The recognition system for online handwritten Bangla characters was available in 2007 by Bhattacharya *et al.* [16]. They have used the direction code based features for recognition and achieved 93.90% accuracy from training sets.

3.3 Recognition of Oriya Script

The Oriya OCR system was developed at Indian Statistical institute, Kolkatta are similar to the Bangla OCR system by the Pal and Chaudhuri [17]. They have used the Hough transform based technique for skew angle estimating and recognizing the Oriya alphabets. In 2007, Pal *et al.* [3] have proposed the recognition system for offline Oriya handwritten script. They have used the curvature feature for this proposal and got the accuracy of about 94.60% from few offline handwritten Oriya samples.

3.4 Recognition of Kannada Script

Kannada is one of the major and earliest script of Southern India and spoken by about more than 50 million people in the Indian state of Andhra Pradesh, Karnataka, Maharashtra and Tamil Nadu. Little work has been done in Kannada handwritten script recognition. The handwritten Kannada numeral recognition is reported by Acharya *et al.* [25]. They have used the structural features and multilevel classifiers for recognition.

3.5 Recognition of Malayalam Script

Malayalam script is also one of the script of Southern India and it is the eighth most popular script in India and spoken by about 30 million people in the Indian state of Kerala. The writing nature of Malayalam is similar to Tamil. Rajashekararadhya and Ranjan [26] have proposed the algorithm for feature extraction of Malayalam script recognition. This algorithm can be used for other Southern Indian scripts like Kannada, Telegu and Tamil also.

3.6 Recognition of Tamil Script

Sundaram and Ramakrishnan [18] have proposed the two dimensional principal component analysis (2DPCA) technique for recognition of online Tamil character recognition. The on-line Tamil character recognition is reported by Aparna *et al.* [19]. They have used shape based features including dot, line terminal, bumps and cusp.

3.7 Recognition of Gurmukhi Script

Gurmukhi script is the script used for writing Punjabi language and is derived from the old Punjabi term “Guramukhi”, which means “from the mouth of the Guru”. Gurmukhi script has three vowel bearers, thirty two consonants, six additional

consonants, nine vowel modifiers, three auxiliary signs and three half characters. Gurmukhi script is 14th most widely used script in the world. Writing style of Gurmukhi script is from top to bottom and left to right. In Gurmukhi script, there is no case sensitivity. Presently, fairly printed Gurmukhi script documents and degraded printed Gurmukhi script documents can be recognized by OCR software, but there are very limited efforts in the recognition of complete handwritten Gurmukhi script document. Most of the work on Gurmukhi script recognition system is done by Lehal and Singh [20]. They have developed the complete recognition system for printed Gurmukhi script, where connected components are first segmented using thinning based approach. Algorithm for segmentation of isolated handwritten words was available in 2006 by Sharma and Lehal [21]. They proposed technique for segments the words in an iterative manner by focusing on presence of headline, aspect ratio of characters and vertical and horizontal projection profiles.

Jindal *et al.* [22] have provided a solution for touching character segmentation of printed Gurmukhi script. Also they have provided a very useful solution for segmenting overlapping lines in various Indian scripts [23]. They have proposed the technique for segment the degraded Gurmukhi script into upper, middle, lower zones. They have also provided the complete recognition system for complete degraded printed Gurmukhi script documents [22, 23]. Online handwriting Gurmukhi script recognition system was available in 2008 by Sharma *et al.* [24]. They have used the elastic matching technique in which character is recognized into two stages. In the first stage they recognize the strokes, in the second stage, character is evaluated on the basis of recognized strokes.

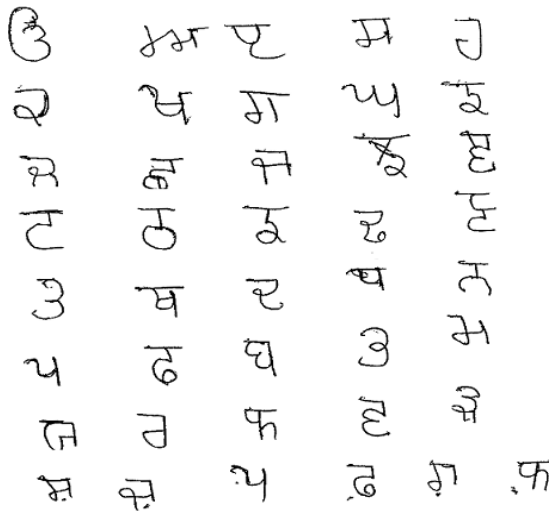


Fig. 6. Handwritten Gurmukhi characters

4 Conclusion and Future Scope

In this paper, we presented work done on handwritten Indian scripts. Firstly, we discussed stages of complete OCR system for handwritten document recognition.

After that we explore the techniques and methods, developed for recognition of particular Indian script. A lot of work has been done on Devanagari and Bangla handwritten characters recognition. Little work has been done to recognize the Oriya, Kannada, Tamil, Malayalam and Gurmukhi handwritten script. But till now there is no complete recognition system is available for recognition of Indian scripts. As such, there is a need for a handwriting OCR for Indian script that can help the people for converting the handwritten text to computer processable format. In future work, the techniques used for Bangla and Devanagari characters can be used for other offline handwritten Indian scripts so that accuracy of recognition can be perfect as Bangla and Devanagari character.

References

1. Pal, U., Chaudhuri, B.B.: Indian Script character recognition: a survey. *Pattern Recognition* 37, 1887–1899 (2004)
2. Pal, U., Chaudhuri, B.B.: Script line separation from Indian multi –script documents. In: *The Proceedings of 5th ICDAR*, pp. 406–409 (1999)
3. Pal, U., Wakabayashi, T., Kimura, F.: A system for off-line Oriya handwritten character recognition using curvature feature. In: *The Proceedings of 10th ICIT*, pp. 227–229 (2007)
4. Chaudhuri, B.B., Pal, U.: An OCR System to Read Two Indian Language Scripts: Bangla and Devanagari (Hindi). In: *The Proceedings of 4th ICDAR*, vol. 2, pp. 1011–1015 (1997)
5. Bansal, V., Sinha, R.M.K.: Integrating Knowledge Sources in Devanagari Text Recognition. *IEEE Transaction on Systems, Man and Cybernetics* 30(4), 500–505 (2000)
6. Mori, S., Yamamoto, K., Suen, C.Y.: Historical review of OCR research and development. *Proceedings of the IEEE* 80(7), 1029–1058 (1992)
7. Kumar, D.: AI approach to hand written Devanagari script recognition. In: *The Proceedings of IEEE Region 10th International Conference on EC3-Energy, Computer, Communication and Control Systems*, vol. 2, pp. 229–237 (2008)
8. Hanmandlu, M., Murthy, O.V.R., Madasu, V.K.: Fuzzy model based recognition of handwritten Hindi characters. In: *The Proceedings of 9th Biennial Conference of the Australian Pattern Recognition Society*, pp. 454–461 (2007)
9. Sethi, K., Chatterjee, B.: Machine recognition of constrained hand-printed Devanagari numerals. *J. Inst. Elec. Telecom. Engg.* 22, 532–535 (1976)
10. Dey, L., Bajaj, R., Chaudhuri, S.: Devanagari numeral recognition by combining decision of multiple connectionist classifier. *Sadhana* 27, 59–72 (2002)
11. Chaudhuri, B.B., Majumder, D.D., Parui, S.K.: A procedure for recognition of connected hand written numerals. *Int. J. Systems Sci.* 13, 1019–1029 (1982)
12. Pal, U., Roy, K., Kimura, F.: Bangla Handwritten Pin Code String Recognition for Indian Postal Automation. In: *The Proceedings of 11th ICFHR*, pp. 290–295 (2008)
13. Bishnu, A., Chaudhuri, B.B.: Segmentation of Bangla handwritten text into characters by recursive contour following. In: *The Proceedings of 5th ICDAR*, pp. 402–405 (1999)
14. Dutta, A., Chaudhuri, S.: Bengali alpha-numeric character recognition using curvature features. *Pattern Recognition* 26(12), 1757–1770 (1993)
15. Pal, U., Belaid, A., Choisy, C.: Touching numeral segmentation using water reservoir concept. *Elsevier Science Inc.* 24(1), 261–272 (2003)
16. Bhattacharya, U., Gupta, B.K., Parui, S.K.: Direction code based features for recognition of online handwritten characters of Bangla. In: *The Proceedings of 9th ICDAR*, vol. 1, pp. 58–62 (2007)

17. Pal, U., Chaudhuri, B.B.: Skew Angle Detection of Digitized Indian Script Documents. *IEEE Transactions on PAMI* 19(2), 182–186 (1997)
18. Sundaram, S., Ramakrishnan, A.G.: Two Dimensional Principal Component Analysis for Online Tamil Character Recognition. In: *The Proceedings of 11th ICFHR*, pp. 88–94 (2008)
19. Aparna, K.H., Subramaniam, V., Kasirajan, M., Prakash, G.V., Chakravarthy, V.S., Madhvanath, S.: Online handwriting recognition for Tamil. In: *The Proceedings of 9th International Workshop on Frontiers in Handwriting Recognition (IWFHR)*, pp. 438–443 (2004)
20. Lehal, G.S., Singh, C.: A Gurmukhi Script Recognition System. In: *The Proceedings of 15th International Conference on Pattern Recognition*, vol. (2), pp. 557–560 (2000)
21. Sharma, D.V., Lehal, G.S.: An iterative algorithm for segmentation of isolated handwritten words in Gurmukhi script. In: *The Proceedings of 18th ICPR*, vol. 2, pp. 1022–1025 (2006)
22. Jindal, M.K., Lehal, G.S., Sharma, R.K.: Segmentation Problems and Solutions in Printed Degraded Gurmukhi Script. *International Journal of Signal Processing* 2(4), 258–267 (2005)
23. Jindal, M.K., Sharma, R.K., Lehal, G.S.: Segmentation of Horizontally Overlapping Lines in Printed Indian Scripts. *International Journal of Computational Intelligence Research* 3(4), 277–286 (2007)
24. Sharma, A., Kumar, R., Sharma, R.K.: Online handwritten Gurmukhi character recognition using elastic matching. In: *The Proceedings of Congress on Image and Signal Processing*, vol. 2, pp. 391–396 (2008)
25. Acharya, D., Subba Reddy, N.V., Makkithaya, K.: Multilevel Classifiers in Recognition of Handwritten Kannada Numerals. In: *The Proceedings of PWASET*, vol. 2, pp. 284–289 (2008)
26. Rajashekararadhya, S.V., Ranjan, P.V.: Efficient Zone Based Feature Extraction Algorithm for Handwritten Numeral Recognition of Four Popular South Indian Scripts. *Journal of Theoretical and Applied Information Technology* 4(12), 1171–1181 (2008)

Satellite Image Processing Using Discrete Cosine Transform and Singular Value Decomposition

B.K. Ashish, A. Kumar, and P.K. Padhy

¹ PDPM Indian Institute of Information Technology Design and Manufacturing,
Jabalpur-482005, MP India
{aryanbhandari08, anilkdee, prabin16}@gmail.com

Abstract. In this paper, a novel contrast enhancement technique for contrast enhancement of a low-contrast satellite image has been proposed based on the singular value decomposition (SVD) and discrete cosine transform (DCT). The singular value matrix represents the intensity information of the given image and any change on the singular values change the intensity of the input image. The proposed technique converts the image into the SVD-DCT domain and after normalizing the singular value matrix; the enhanced image is reconstructed by using inverse DCT. The visual and quantitative results suggest that the proposed SVD-DCT method clearly shows the increased efficiency and flexibility of the proposed method over the exiting methods such as the histogram equalization, gamma correction and SVD-DWT based techniques.

Keywords: Singular Value Decomposition (SVD), discrete cosine transforms (DCT), image equalization and satellite image contrast enhancement.

1 Introduction

Image enhancement is the technique which is most widely required in the field of image processing to improve the visualization of the features [1]. Satellite images are useful in many applications for instance finding of the spectral signature of different objects such as the vegetation index, land cover classification, and crop pattern. The remote sensing image is vital in the areas of unfortunate natural disasters to provide humanitarian aid and damage assessment as well as to devise new protection strategies [2]. One of the most common problem, occurs in satellite images while capturing image with a huge amount of distance, is the dark light and contrast of image. Contrast is determined by the difference in the color and brightness of the object with other objects in the same field of view [2]. Basically contrast is developed due to difference in luminance which is reflected from two surfaces.

If an image has been taken in very dark or a very bright situation, the information may be lost in those areas which are excessively and uniformly dark or bright [2]. The problem is how the contrast of an image can be improved from the input satellite image which has complete information but is not visible. There have been several technique reported in literature for the contrast analysis of satellite image such as General Histogram Equalization (GHE), Gamma correction and local histogram

equalization (LHE) [2, 3]. These techniques are very simple and effective Indies for the contrast enhancement [4]. But these techniques are not efficient as the information laid on the histogram of the image which is totally lost.

During the last decade, the Wavelet Transform, more particularly Discrete Wavelet Transform has emerged as powerful and robust tool for analyzing and extracting information from non-stationary signal such as speech signals due to the time varying nature of these signals. Discrete wavelet transform is also widely useful in many fields like removing of noise, compression, enhancement and remote sensing [3]. Recently, many techniques [4] have been developed based on wavelet or wavelet packets for analysis of satellite images.

In this paper, a novel technique based on the singular value decomposition (SVD) and discrete cosine transform (DCT) has been proposed for enhancement of low-contrast satellite images. SVD technique is based on a theorem from linear algebra which says that a rectangular matrix A, that can be broken down into the product of three matrices, as follows: (i) an orthogonal matrix U_A , (ii) a diagonal matrix Σ_A and (iii) the transpose of an orthogonal matrix V_A [4]. The singular-value-based image equalization (SVE) technique is based on equalizing the singular value matrix obtained by singular value decomposition (SVD) [1, 4, 5]. SVD of an image, which can be interpreted as a matrix, is written as follows:

$$A = U_A \Sigma_A V_A^T \tag{1}$$

where U_A and V_A are orthogonal square matrices known as hanger and aligner, respectively, and the Σ_A matrix contains the sorted singular values on its main diagonal and basic enhancement occurs due to scaling of singular values of the DCT coefficients [4, 5]. The singular value matrix represents the intensity information of image and any alteration on the singular values change the intensity of the input image. The main advantage of using SVD for image equalization comes from the fact that Σ_A contains the intensity information of the image; here multiband satellite image has been taken for better the analysis [1, 5]. In the case of singular value decomposition the ratio of the highest singular value of the generated normalized matrix, with mean zero and variance of one, over a particular image can be calculated by the equation as given below:

$$\xi = \frac{\max(\Sigma_{N(\mu=0, \text{var}=1)})}{\max(\Sigma_A)} \tag{2}$$

where $\Sigma_{N(\mu=0, \text{var}=1)}$ is the singular value matrix of the synthetic intensity matrix. This coefficient can be used to regenerate an equalized image using

$$E_{\text{equalized } A} = U_A (\xi \Sigma_A) V_A^T \tag{3}$$

where $E_{\text{equalized } A}$ is used to denote the equalized image named A. The equalization of an image is used to remove the problem of the illumination, which is basically one cause of the low contrast image and blurring [1, 6].

Discrete cosine transform (DCT) is applied to extract texture features of an image. Texture segmentation is an important task in many applications [7]. It can be combined with other low-level features segmentation to improve the performance of

the feature which is extracted from image [8, 9]. The DCT converts a spatial domain waveform into its constituent frequency components as represented by a set of coefficients. The process of reconstructing a set of spatial domain samples is called the Inverse Discrete Cosine Transform (IDCT) [10]. DCT is used to separate the lower and higher frequency coefficient in two parts. The DCT features are stored in the form of an array where the required features of the image can be calculated by DCT frequency coefficients [11]. Using the initial few values, it delivers reasonable performance for feature extraction of the satellite images [10, 11]. Hence, after inverse DCT (IDCT), the enhanced image will be more effective, sharper and having a good contrast [12].

In above context, therefore, in this paper a new technique is proposed based on combined effect of SVD and DCT.

2 Overview of DCT and SVD

DCT was first time used in 1974 [7]. The DCT coefficients can be quantized using visually-weighted quantization values. DCT is a fast algorithm similar to FFT [13, 14].

2.1 DCT

The discrete cosine transform is a technique for converting a signal into elementary frequency components. It is widely used for extracting the features [14]. The one-dimensional DCT is useful in processing of one-dimensional signals such as speech waveforms. For analysis of the two-dimensional (2-D) signals such as images, a 2-D version of the DCT is required. The DCT works by separating images into parts of differing frequencies [14, 15].

For an $N \times M$ matrix, the 2D-DCT is computed in a simple way. Initially, 1D-DCT is applied to each row of the matrix and then, to each column of the matrix 'x'. Thus, the transform of x is given by

$$y(u, v) = \sqrt{\frac{2}{M}} \sqrt{\frac{2}{N}} \alpha_u \alpha_v \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} x(m, n) \cos \frac{(2m+1)u\pi}{2M} \cos \frac{(2n+1)v\pi}{2N} \tag{4}$$

where

$$\alpha_v = \left\{ \begin{array}{ll} \frac{1}{\sqrt{2}} & v = 0 \\ 1 & v = 1, 2, \dots, N-1 \end{array} \right\} \quad \& \quad \alpha_u = \left\{ \begin{array}{ll} \frac{1}{\sqrt{2}} & u = 0 \\ 1 & u = 1, 2, \dots, N-1 \end{array} \right\}$$

The image is reconstructed by applying inverse DCT operation according to Eq. 5:

$$x(m, n) = \sqrt{\frac{2}{M}} \sqrt{\frac{2}{N}} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} \alpha_u \alpha_v y(u, v) \cos \frac{(2m+1)u\pi}{2M} \cos \frac{(2n+1)v\pi}{2N} \tag{5}$$

Since, the 2D-DCT can be computed by applying 1D transforms separately to the rows and columns; it means that the 2D-DCT is separable in the two dimensions. As

in the one-dimensional, each element $y(u,v)$ of the transform is the inner product of the input and a basis function, but in this case, the basic functions are $M \times N$ matrices [16]. $x(m, n)$ is the x, y^{th} element of the image represented by the matrix y , M is the size of the block of image on which the DCT is applied. Eqn. (5) calculates on entry (u, v^{th}) of the transformed image from the pixel values of the original image matrix [17, 18].

The DCT helps to separate the image into parts (or spectral sub-bands) of differing importance (with respect to the image's visual quality). The DCT is similar to the discrete Fourier transform: it transforms a signal or image from the spatial domain to the frequency domain as show in Fig.1. The popular block-based DCT transform segments an image non-overlapping block and applies DCT to each block. It gives result in three frequency sub-bands: low frequency sub-band, mid-frequency sub-band and high frequency sub-band. DCT-based enhancement is based on two facts. The first fact is that much of the signal energy lies at low-frequencies sub-band which contains the most important visual parts of the image. The second fact is that high frequency components of the image [17, 18]. The basic operation of the DCT is as follows: the input multiband satellite image is M by N , $f(i, j)$ is the intensity of the pixel in rows i and column j , $s(x, y)$ is the DCT coefficient in row k_1 and column k_2 of the DCT matrix. For most multiband satellite images, much of the signal energy lies at low frequencies; these appear in the upper left corner of the DCT.

2.2 SVD

Singular value decomposition (SVD) can be calculated mainly by the three mutually compatible points of view [18]. On the one hand, we can view it as a method for transforming correlated variables into a set of uncorrelated ones that better expose the various relationships among the original data items [19]. At the same time, SVD is a method for identifying and ordering the dimensions along which data points exhibit the most variation. This ties in to the third way of viewing SVD, which is that once we have identified where the most variation is present, then it is possible to find the best approximation of the original data points using fewer dimensions [20]. Hence, SVD can be seen as a method for data reduction and mostly for feature extraction as well as for the enhancement of the low contrast images [21]. Following are the basic ideas behind SVD: taking a high dimensional, highly variable set of data points and reducing it to a lower dimensional space that exposes the substructure of the original data more clearly and orders it from most variation to the least [22].

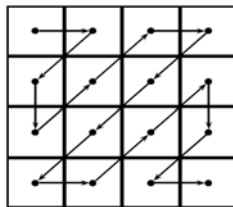


Fig. 1. Zigzag (highest to lowest energy coefficients of DCT)

3 Proposed Methodology

In this paper basically two parts involve in the enhancement of the satellite image. The first one is SVD, as it is mentioned in Section 2 and the singular value matrix [22]. The SVD contains the illumination information in the image. So that, converting the singular values will directly change the illumination of the image, therefore the other information present in the image will be as same as before [23]. The second most significant aspect of this method is the application of DCT as it is discussed in previous section.

The DCT can be regarded as a discrete-time version of the Fourier-Cosine series [5-7]. It is a close relative of DFT, a technique for converting a signal into elementary frequency components [23, 24]. Thus DCT can be computed with a Fast Fourier Transform (FFT). Unlike DFT, DCT is real-valued and provides a better approximation of a signal with few coefficients [7-9]. This approach reduces the size of the normal equations by discarding higher frequency DCT coefficients, while avoiding the overhead of image resampling in the DCT domain [9-11]. In the DCT domain, each significance equation is dependent on the corresponding DCT frequency, and the frequency characteristics of the important structural information in the image. For the aerial images in our study, the important structural information is present in the low frequency DCT coefficients [12]. Hence, separating the high-frequency DCT coefficient and applying the illumination enhancement in the low-frequency DCT coefficient, it will collect and cover the edge information from possible degradation of the multiband remote sensing satellite images [13-15]. DCT output is followed by applying IDCT, and thereby reconstructing the final image by using IDCT, the consequence image will be enhanced with respect to illumination and it will be sharper with good contrast [24].

In the proposed technique, initially the input satellite image ‘A’ for processed by GHE to generate \hat{A} . After getting this, both of these images are transformed by DCT into the lower-frequency DCT coefficient, and higher-frequency DCT coefficient [15-18]. Then, the correction coefficient for the singular value matrix can be calculated by using

$$\xi = \frac{\max(\Sigma_{D^{\wedge}})}{\max(\Sigma_D)} \tag{6}$$

where $(\Sigma_{D^{\wedge}})$ is the lower-frequency coefficient singular matrix of the satellite input image, and (Σ_D) is the lower-frequency coefficient singular matrix of the satellite output image of the General Histogram Equalization (GHE). The new satellite image (D) is examined by

$$\bar{\Sigma}_{D^{\wedge}} = \xi \Sigma_D \tag{7}$$

$$\bar{D} = U_D \bar{\Sigma}_{D^{\wedge}} V_D \tag{8}$$

Now, \bar{D} , is the lower DCT frequency component of the original image that is reorganized by applying the inverse operation (IDCT) to produce the consequence equalized image \bar{A} as given by Eqn. (8)

$$\bar{A} = IDCT(\bar{D}) \tag{9}$$

Following steps are to be undertaken to discuss the main computational process of the proposed algorithm:

Step1: In the very first step, a low contrast input satellite image has been taken for the analysis.

Step2: Equalize the satellite image using general histogram equalization technique.

Step3: After equalization, compute the discrete cosine transform for the contrast enhancement.

Step4: In this step calculate the two variables D^\wedge and D from the discrete cosine transform image.

Step5: After getting D^\wedge and D , SVD is applied for the calculation of the U, Σ, V and find the max element in Σ .

Step6: Calculate $\text{Max } \Sigma_D$ & $\text{Max } \Sigma_{D^\wedge}$ with the help of singular value decomposition process.

Step7: Calculate ξ using Eqn. (6) $\xi = \text{max}(\Sigma_{D^\wedge}) / \text{max}(\Sigma_D)$

Step8: Calculate the new Σ_{D^\wedge} using equation (7) and (8) $\Sigma_{D^\wedge} = \xi \Sigma_D$ & $\bar{D} = U_D \bar{\Sigma}_{D^\wedge} V_D$

Step9: Apply IDCT after getting new $\bar{\Sigma}_{D^\wedge}$

Step10: Equalized satellite image.

4 Results and Discussion

In this section, the proposed method has been used for enhancement of the several satellite images. Different satellite images are included to demonstrate the usefulness of this algorithm. The performance of this method is measured in terms of following significant parameters:

$$\text{Mean } (\mu) = \frac{1}{MN} \sum_{x=1}^{M-1} \sum_{y=1}^{N-1} I(x-y) \tag{10}$$

$$\text{Standard Deviation } (\sigma) = \sqrt{\frac{1}{MN} \sum_{x=1}^{M-1} \sum_{y=1}^{N-1} \{I(x,y) - \mu\}^2} \tag{11}$$

Mean (μ) is the average of all intensity value. It denotes average brightness of the image, where as standard deviation is the deviation of the intensity values about mean. It denotes average contrast of the image. Here $I(x, y)$ is the intensity value of the pixel (x, y) , and (M, N) are the dimension of the image.

A flowchart of the proposed methodology is illustrated in Fig. 2.

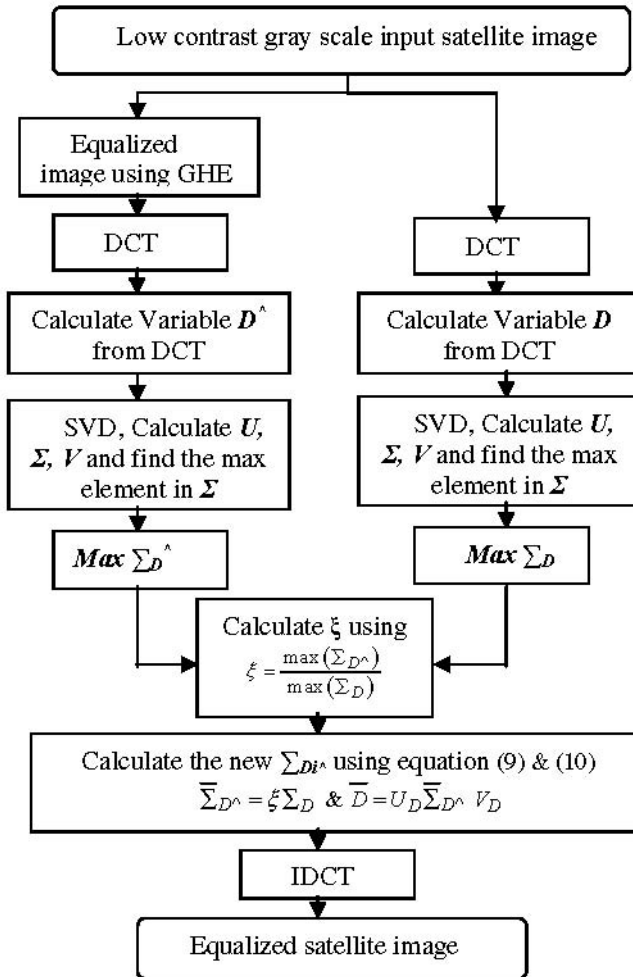


Fig. 2. Flow chart of the proposed mythology

First of all, the performance of the proposed algorithm is carried out in the different images sample. Thereafter, the comparison of proposed method is done with the following method: (1). Gamma correction method, (2). GHE and (3). DWT technique and it show the superiority of the proposed technique.

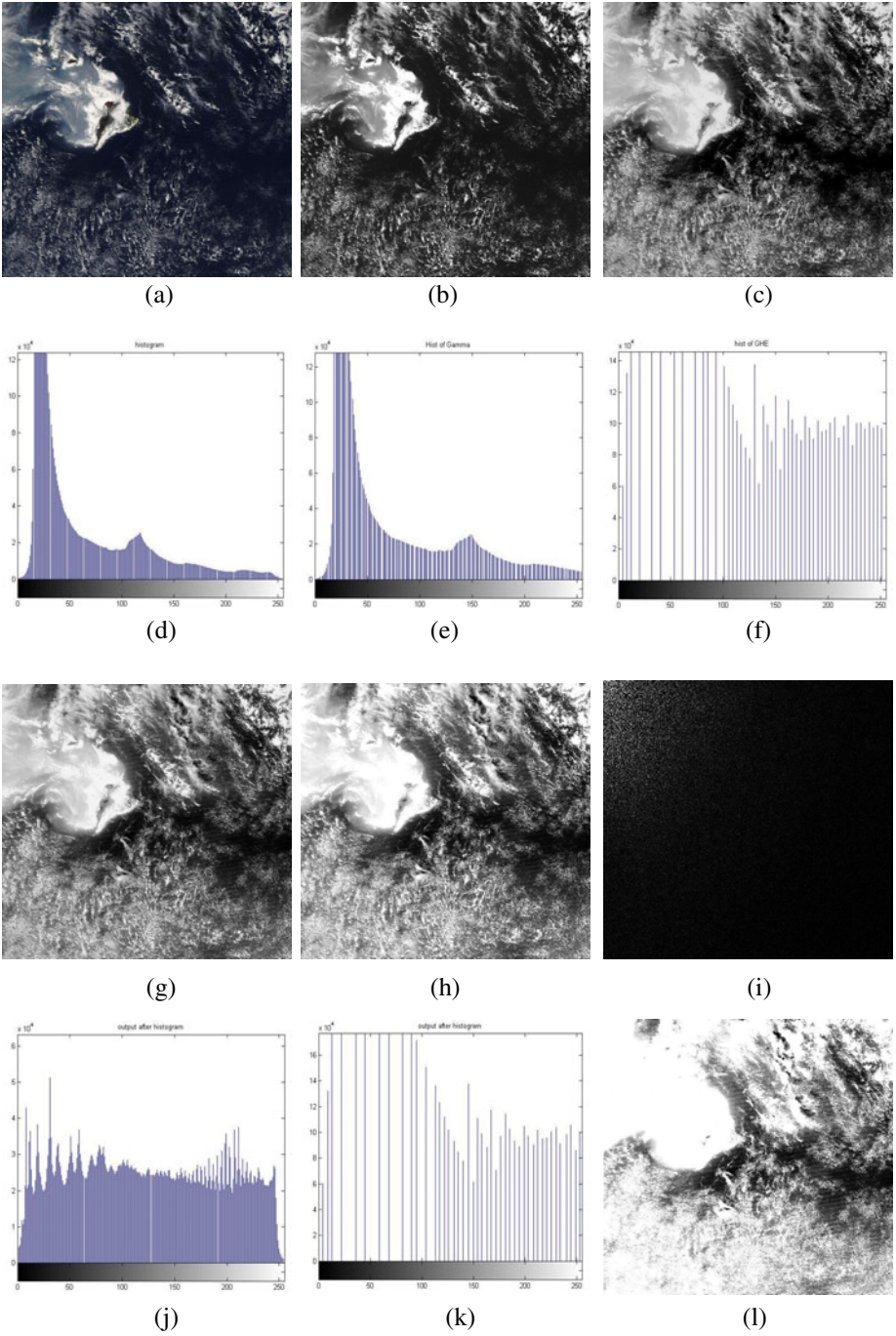


Fig. 3. Various resultant images using existing techniques and proposed technique

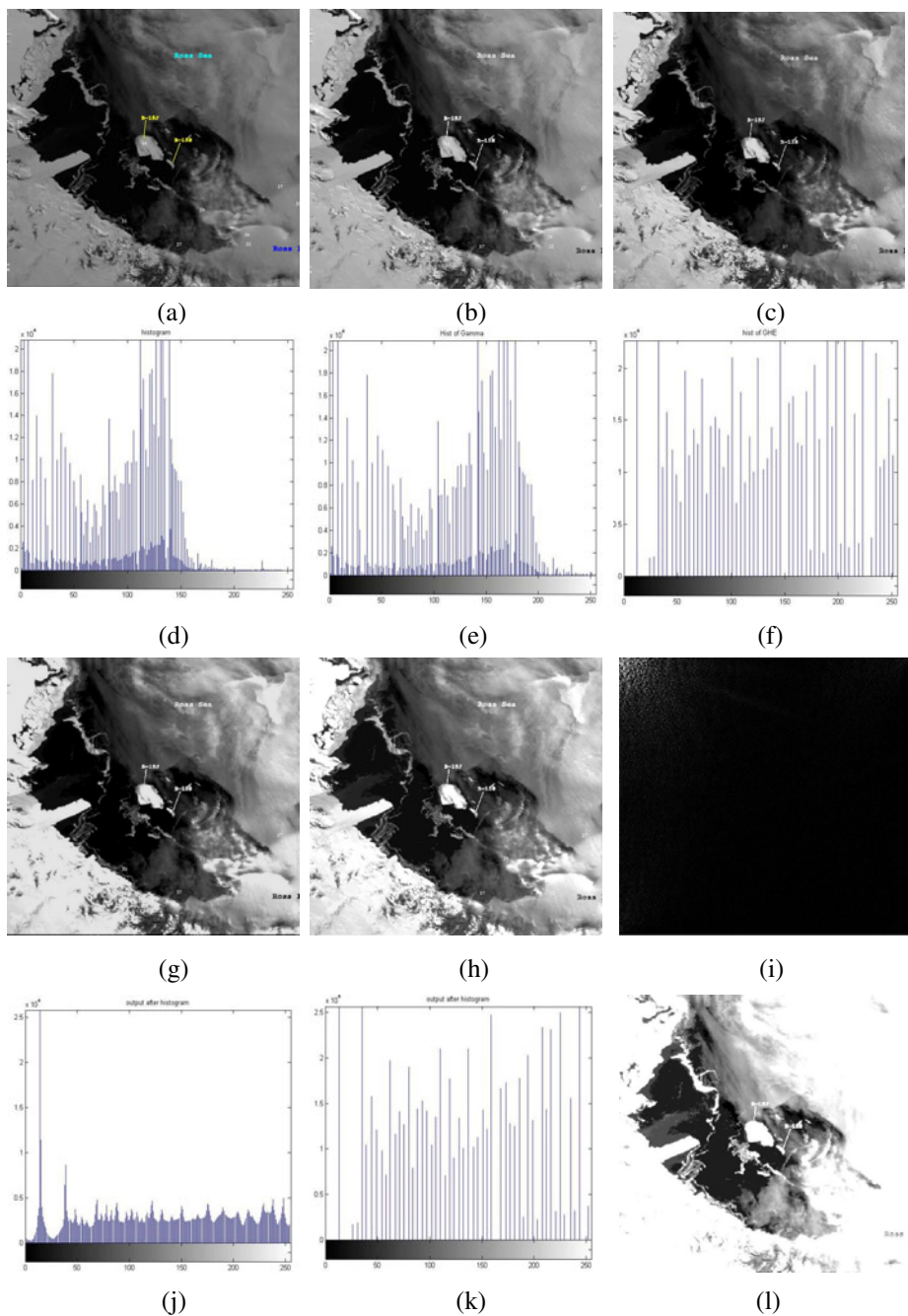


Fig. 4. Various resultant images using existing techniques and proposed technique

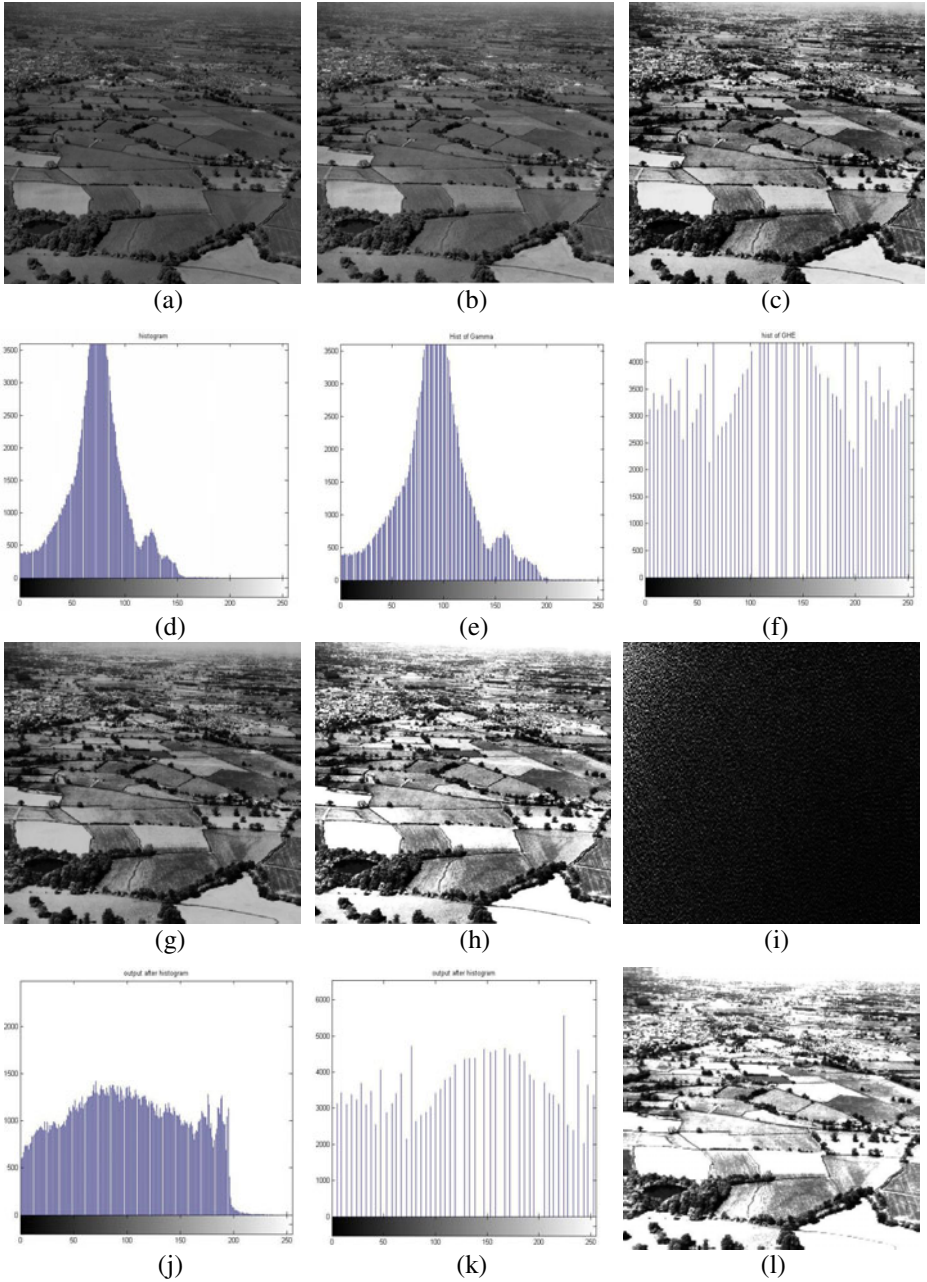


Fig. 5. Various resultant images using existing techniques and proposed technique

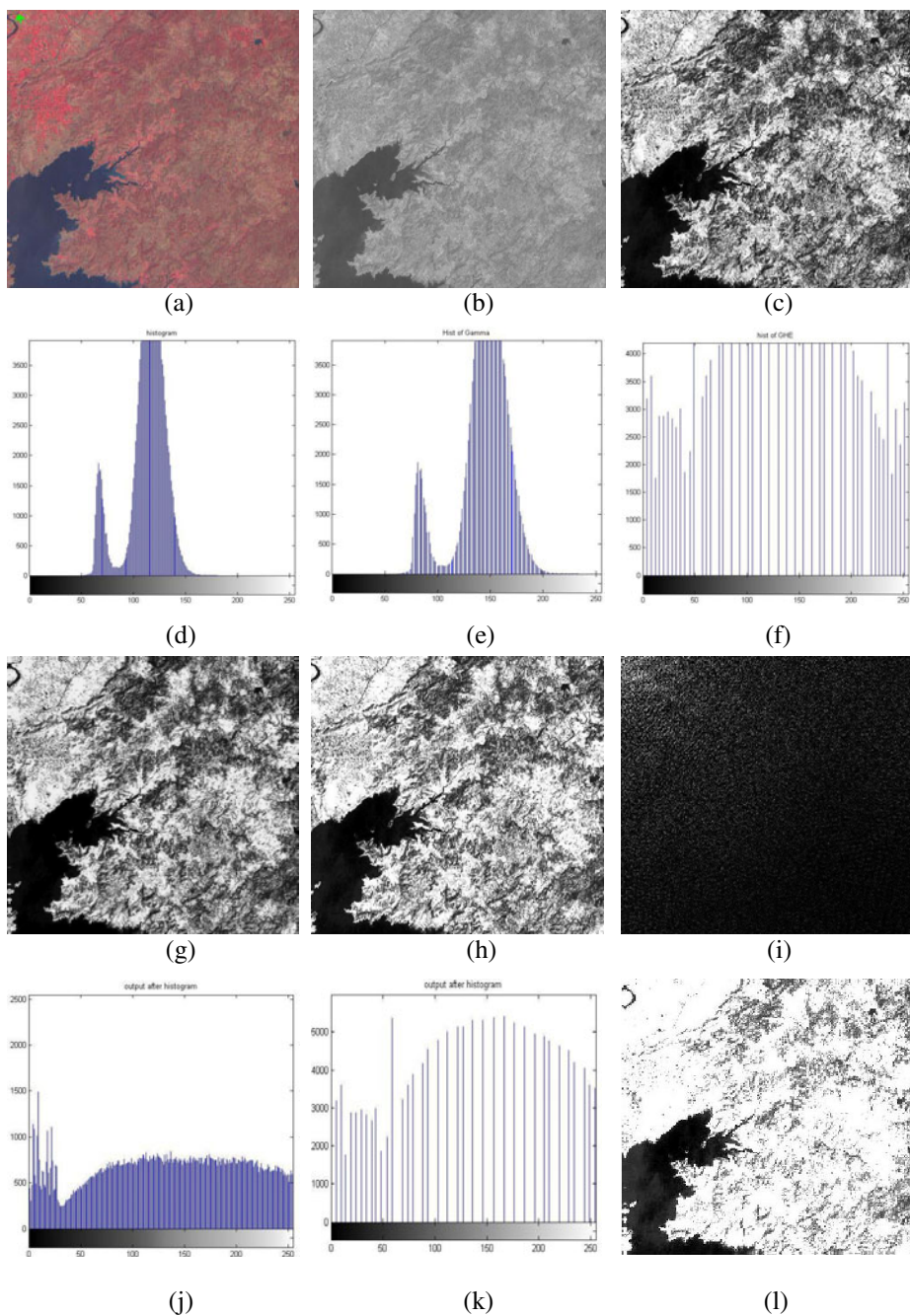


Fig. 6. Various resultant images using existing techniques and proposed technique

[Fig.3, Fig.4, Fig.5 and Fig.6]: (a) Low contrast original image for analysis, (b) Equalized image by using Gamma Correction technique, (c) Equalized image by the GHE, (d) Histogram of the input satellite image, (e) Histogram of the Gamma Correction output image (f) Histogram of the General Histogram Equalization output image, (g) Equalized image by the SVD-DWT technique, (h) Equalized image by the proposed methodology using SVD-DCT technique, (i) Lower and higher frequency coefficient after applying DCT, (j) Histogram of the SVD-DWT output image, (k) Histogram of the proposed methodology using SVD-DCT technique and (l) LL frequency component using DWT technique.

Here, four different satellite images of different region have been taken for enhancement. The simulation results obtained with the proposed methodology and other existing methods (Gamma Correction, General Histogram Equalizer and DWT) are included in Table 1 and graphically shown in the figures (3), (4), (5) and (6). Fig.6. shows the low-contrast image of Jabalpur region and the corresponding enhanced images. The performance parameters in each method are:

Gamma Correction: $\mu=75.3408$ $\sigma=4.9394e+003$;

General Histogram Equalizer: $\mu=127.4132$ $\sigma=5.6228e+003$;

Discrete wavelet Transform $\mu=122.9445$ $\sigma=4.9866e+003$;

Proposed Methodology: $\mu=142.2035$ $\sigma=7.0040e+003$;

Table 1. Comparison of the results between different proposed Methodology and already existing technique

Sr. No	Input Image Mean (μ) & Standard Deviation (σ)	Gamma Correction Image Mean (μ) & Standard Deviation (σ)	GHE Image Mean (μ) & Standard Deviation (σ)	Comparison with existing DWT technique (μ) & (σ)	Proposed IDCT Image Mean (μ) & Standard Deviation (σ)
1	$\mu=62.9134$ $\sigma=2.8248e+003$	$\mu=75.3408$ $\sigma=4.9394e+003$	$\mu=127.4132$ $\sigma=5.6228e+003$	$\mu=122.9445$ $\sigma=4.9866e+003$	$\mu=142.2035$ $\sigma=7.0040e+003$
2	$\mu=85.0754$ $\sigma=2.6100e+003$	$\mu=107.5585$ $\sigma=4.3429e+003$	$\mu=127.1981$ $\sigma=5.5794e+003$	$\mu=133.8994$ $\sigma=7.9961e+003$	$\mu=138.9535$ $\sigma=6.6583e+003$
3	$\mu=72.7006$ $\sigma=787.3876$	$\mu=90.4819$ $\sigma=1.3150e+003$	$\mu=127.5886$ $\sigma=5.6067e+003$	$\mu=97.8396$ $\sigma=2.9958e+003$	$\mu=150.2097$ $\sigma=7.7711e+003$
4	$\mu=114.8540$ $\sigma=937.7782$	$\mu=142.7744$ $\sigma=626.8973$	$\mu=127.6460$ $\sigma=5.5899e+003$	$\mu=141.0827$ $\sigma=5.8780e+003$	$\mu=154.1794$ $\sigma=8.1553e+003$

The quality of the visual results indicates that the proposed technique is sharper and brighter than existing technique as compared. After obtaining mean and standard deviation, it is found that the proposed algorithm gives better results in comparison with the existing techniques. Mean (μ) represent the intensity of the image and the

standard deviation represent (σ) the contrast present in the images. The proposed DCT method represents the better contrast as well as better brightness with appropriate contrast. However, the estimated mean (μ) and standard deviation (σ) in **Figs.** [3(e), 4(e), 5(e) and 6 (e)] of the proposed method covers a good range of gray level and this is the cause of the better illumination. Therefore the observation of the proposed DCT gives the better result.

In order to exhibit the superiority of the proposed methodology three different images have been taken for analysis. The singular values denote luminance of each image layer after decomposition using DCT methodology.

After analysis of all the technique, it has been found that the proposed technique indicates the better mean (μ) and standard deviation (σ) in comparison with Gamma Correction, GHE and DWT. The quality of the input image was poor but after applying the DCT, the result is optimized with reference to brightness and contrast. The histograms obtained from of the proposed technique are stretched in dynamic range, thereby signifying the improvement in contrast of the output image. The mechanism of contrast enhancement can be attributed to scaling of singular values of the DCT coefficients. Since singular values denote luminance of each image layer after decomposition, scaling of these values leads to variation (enhancement) of luminance of each layer and hence leads to overall contrast enhancement. Thus, it is evident that the proposed methodology based on SVD and DCT improve the image quality as well as contrast of the images.

5 Conclusions

In this paper, a novel technique based on SVD-DCT domain for enhancement of low-contrast satellite images has been proposed. The basic enhancement occurs due to scaling of singular values of the DCT coefficients. Performance of this technique has been compared with existing contrast enhancement techniques like histogram equalization, gamma correction and SVD-DWT based techniques. The experimental results show that the proposed technique gives better performance in terms of contrast (variance) as well as brightness (mean) of the enhanced image as compared to the other existing techniques. Thus, this technique can be considered suitable for enhancement of low contrast satellite image.

References

1. Demirel, H., Ozcinar, C., Anbarjafari, G.: Satellite Image Contrast Enhancement Using Discrete Wavelet Transform and Singular Value Decomposition. *IEEE Geosciences and Remote Sensing Letters* 7(2), 333–337 (2010)
2. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*. Prentice-Hall, Englewood Cliffs (2007)
3. Demirel, H., Anbarjafari, G., Jahromi, M.N.S.: Image Equalization Based On Singular Value Decomposition. In: *Proceeding of IEEE Conference on Computer and Information Sciences*, pp. 1–5 (2008)
4. Hemes, G.M., Danaher, S., Murray, A.: Characterization of Forestry Species - A Comparison Using Singular Value Decomposition (SVD) and Artificial Neural Networks (ANN). In: *Proceeding of IEEE Conference on Image Processing and its Applications*, July 4-6, pp. 815–819 (1995)

5. Murty, P.S., Rajesh, K.P.: A Robust Digital Image Watermarking Scheme Using Hybrid DWT-DCT-SVD Technique. *International Journal of Computer Science and Network Security* 10(1), 185–192 (2010)
6. Sverdlovsk, A., Dexter, S., Eskicioglu, A.M.: Robust DCT-SVD Domain Image Watermarking For Copyright Protection: Embedding Data In All Frequencies. In: 13th European Conference on Signal Processing, September 3-5, pp. 1–4 (2005)
7. Reeves, R., Kubik, K.: Benefits of Hybrid DCT Domain Image Matching. In: *International Archives of Photogrammetric and Remote Sensing*, Amsterdam, vol. XXXIII, Part B3, pp. 771–778 (2000)
8. Pun, C.M., Zhu, H.M.: Image Segmentation Using Discrete Cosine Texture Feature. *International Journal of Computers* 4(1), 19–26 (2010)
9. Sagheer, A., Tsuruta, N., Taniguchi, R.I., Maeda, S.: Hyper-Column Model vs. Fast DCT for Feature Extraction in Visual Arabic Speech Recognition. In: *IEEE Proceeding in Signal Processing and Information Technology*, pp. 761–766 (2005)
10. Khaleel, T.A.: Enhancement of Spatial Structure of an Image by Using Texture Feature Extraction. *Al-Rafidain Engineering* 15(1), 27–37 (2007)
11. Su Kim, K., Lee, M.J., Lee, H.K.: Blind Image Watermarking Scheme in DWT-SVD Domain. In: *IEEE Intelligent Information Hiding and Multimedia Signal Processing*, November 26-28, pp. 477–480 (2007)
12. Azam, M., Anjum, M.A., Javed, M.Y.: Discrete Cosine Transform (DCT) Based Face Recognition in Hexagonal Images. In: *Computer and Automation Engineering (ICCAE)*, February 26-28, vol. 2, pp. 474–479 (2010)
13. Christopher, C.J., Prabukumar, M., Baskar, A.: Color Image Enhancement in Compressed DCT Domain. *ICGST - GVIP Journal* 10, 31–38 (2010)
14. Sanderson, C., Paliwal, K.K.: Fast feature extraction method for robust face verification. *Electronics Letters* 5th 38(25) (December 2002)
15. Sorwar, G., Abraham, A.: DCT Based Texture Classification Using Soft Computing Approach
16. Sorwar, G., Abraham, A., Dooley, L.S.: Texture Classification Based on DCT and Soft Computing. In: *IEEE International Conference on Fuzzy Systems 92-5*, vol. 2, pp. 445–448 (2001)
17. Watson, A.B.: Image Compression Using the Discrete Cosine Transform. *Mathematica Journal*, 81–88 (1994)
18. Cabeen, K., Gent, P.: Image Compression and the Discrete Cosine Transform
19. Sverdlov, A., Dexter, S., Eskicioglu, M.: Robust DCT-SVD Domain Image Watermarking For Copyright Protection. In: *Proceedings of the 2004 Workshop on Multimedia and Security* (2004)
20. Gorodetski, V., Popyack, L., Samoilov, V., Skormin, V.: SVD-based Approach to Transparent Embedding Data into Digital Images
21. Henies, G., Selige, T., Danaher, S.: Singular Value Decomposition in Applied Remote Sensing, pp. 5/1–5/6
22. Boardman, J.W.: Inversion of Imaging Spectrometry Data Using Singular Value Decomposition, pp. 269–272
23. Gonzalez, C.R., Woods, R.E., Eddins, S.L.: *Digital Image processing Using MATLAB* Pearson Education. Second Indian Reprint (2005)
24. Haj, A.A.: Combined DWT-DCT Digital Image Watermarking. *Journal of Computer Science* 3(9), 740–746 (2007)

Physical Layer Authentication in Wired and Wireless Communication Systems

Chandra Mouli and Sanjay Singh

Department of Information and Communication Technology
Manipal Institute of Technology, Manipal University, Manipal, India
chandramouli1987@gmail.com, sanjay.singh@manipal.edu

Abstract. Authentication is the process by which it is verified that someone is who they claim they are. Authentication is a security service that checks the identity of the party at the other end of the line. Most of the times, security is provided above physical layer, but in this paper we are concentrating on authentication at the physical layer in wired and wireless communication systems. The authentication signal is generated from modulated signal using shared secret key by transmitter and receiver and these two signals are superimposed as in the spread spectrum technique. This authentication signal is transmitted concurrently with data in physical medium without using extra bandwidth. For recovering the authentication signal at receiver side MMSE equalizer is used to reduce intersymbol interference (ISI) and noise. This authentication is designed to be stealthy and secured. The simulation results have shown the proof of concept in this paper using Matlab7.6.

1 Introduction

Security has become an implicit part for any communication system. As part of providing security we have a set of security services like confidentiality, integrity, authentication etc. In this paper we focus on the authentication where identity of the sender can be verified at the receiver. Mostly we have seen the authentication mechanisms exist above the physical layer, but in this paper we are applying authentication techniques at the physical layer. Physical layer is the lowest layer of the OSI model that connects the two communicating parties via physical link. Digital communication techniques are used in this layer to transmit message in the form of raw bits over the physical medium. The authentication mechanism is applied to the message signal which generates an authentication signal which is transmitted over the physical link. This authentication signal is transmitted concurrently with the data without any extra bandwidth and transmission cost. The transmission is done through wired and wireless channels. This transmission shows characteristics that, the signal is stealthy and secured by taking some test scenarios.

This paper diverges from much of the previous work research in authentication systems and mechanisms have mostly focused above the physical layer. There are two paradigms of adding authentication: multiplexing or embedding. Simmon

in [1] has given an overview of the methods of adding authentication and how it is received with data as same quality. However, data throughput is penalized since some of the bits carry authentication instead of data. Digital watermarking follows the paradigm of embedded signaling by modifying the data in a controlled manner that provides additional information to the receiver. Authentication may be transmitted as discussed in [2][3] and the addition of authentication to data is stealthy. Unlike the multiplexing approach, embedding additional information degrades the data quality [4].

At the physical layer, authentication is done at sender and receiver by using spread spectrum methods, such as Direct Sequence Spread Spectrum (DSSS) and Frequency Hopping Spread Spectrum (FHSS) as in [5]. Here the sender and receiver shared a secret pseudo random sequence, where sender is authenticated if receiver can demodulate the signal this shows the use of low cost bandwidth and allows communication between the authenticated parties. These techniques are covert and provide robustness to interference; they achieve this at the cost of bandwidth expansion and allow only authenticated parties with knowledge of the secret to participate in communication.

Suppose we want to add authentication to a system in a stealthy way so that users are unaware of the authentication can continue to communicate without any modifications to the protocol. Our approach to authentication exists at the physical layer, and may be used together with spread-spectrum methods or other security schemes at the higher layers to provide a more secure system. The idea of transparently adding information at the physical layer has been discussed for some specific cases. Supangkat in [6] has proposed one such authentication scheme for telephony where an encrypted hash of the conversation is added back into the signal. Similarly, Kleider in [7] has proposed a scheme where a low-power watermark signal is added to the data signal with spread-spectrum techniques.

Paul L. Yu et al. in [8] introduces authentication at physical layer and transmits the authentication signal concurrently along with data without any extra bandwidth or transmission power. The authors have used Rayleigh channel for simulating physical layer authentication. Further they have shown their method of authentication is stealthy and secure.

This paper is organized as follows. Section 2 discusses various aspects of the proposed system. Simulation results and discussion is given in section 3. Finally section 4 concludes this paper.

2 Proposed Scheme

At physical layer, authentication and tag signal are generated from the message signal and these two signals are then superimposed to obtain tagged signal. This tagged signal and the reference signal is transmitted over the physical link to receiver and verify the integrity and authenticity of the message signal.

In this paper, we consider the basic scenario as depicted in Fig.1 where we have sender Alice, receiver Bob and interceptor Eve. Before transmitting an authentication signal from Alice to Bob, they both agree on keyed authentication

method which uses a shared secret key. This allows Bob to verify the integrity and authenticity of received signal by using the shared key. Eve might know, the authentication mechanism but she doesn't know the shared secret key, so here Bob and Eve are conscious receivers but Eve is an adversary receiver, so the signal should be stealthy that is, it should not be easily detectable.

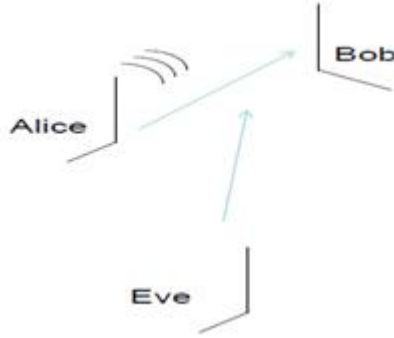


Fig. 1. Basic Communication Scenario

Authentication mechanism are prone to attacks such as, Eve intercepts the signal and changes the reference or authentic signal and sends it to the Bob. Then Bob verifies the signal, if signal is not authentic then the signal will be rejected.

The complete design of proposed scheme is as shown in Fig.2. We have considered two communication channel wired and wireless respectively to transmit authentication signal. Input message is modulated using modulation techniques such as BPSK, QAM etc, as in [5]. The modulated signal is passed through the

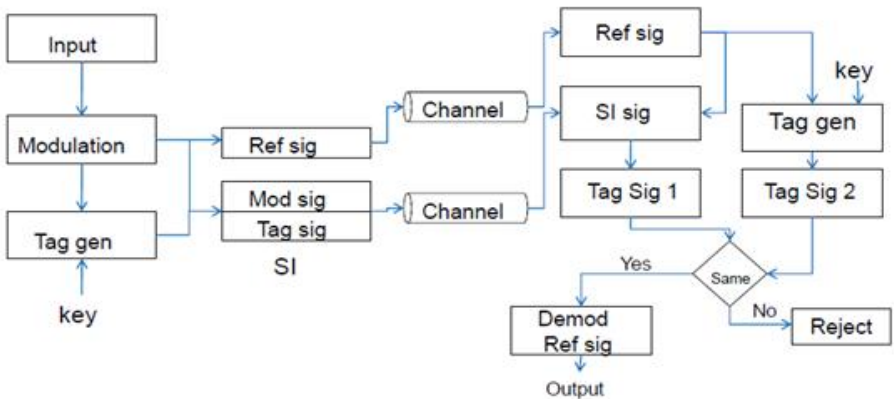


Fig. 2. Block Diagram of the Proposed System

tag generator to generate a tag signal as explained in the next section. This tag signal is superimposed with the modulated signal which is called superimposed (SI) signal. This SI signal and reference signal sent through the communication channel. At the receiver, SI signal and the reference signals are received. From the received reference signal, a tag signal is generated, another tag signal is extracted from SI and reference signals and both the signals are compared. If the two signals differ then received signal is not authentic and it is rejected. If the received tag signal and the regenerated tag at the receiver matches then the received signal is authentic and it is demodulated.

2.1 Transmitter

Let the messages are blocks of N symbols denoted by

$$b = \{b_1, b_2, \dots, b_n\}.$$

These message symbols assumed as an i.i.d (Independent, identically distributed) random variables. This message symbol is to be modulated using any of the modulation technique such as QAM, BPSK etc. The message signal is given by

$$s = f_m(b) \quad (1)$$

where $f_m(\cdot)$ is the modulation function. This modulated signal is passed through the tag generator.

2.2 Tag Generator

In tag generator, keyed authentication scheme [2] is used as shown in Fig.3. The modulated signal is in complex form but tag generator takes input in binary form. To convert modulated signal from complex to binary form we use RCBNS (Redundant Complex Binary Number System) method as stated in [9]. The Redundant Complex Binary Number System (RCBNS) is developed by combining a Redundant Binary Number and a complex number in base $(-1+j)$. A Redundant Complex Binary Number System consists of both real and imaginary-radix number systems that form a redundant integer digit set. The pseudo code of this RCBNS method is given in Algorithm. [1].

After converting modulated signal into binary form it is passed through keyed hash function (HMAC) [10] which is also called MAC's (Message Authentication code) which is used to verify the integrity of a message. Here we have used SHA512 as hash function to generate hash using HMAC, but MD5, SHA1, SHA384 etc can also be used to generate a hash. There is a shared secret key between the Alice and Bob. This secret key is encrypted with 128-bit AES algorithm [10] and given to the tag generator function. This tag signal is in the form of binary bit stream. The tag signal can be represented as:

$$t = g(s, k) \quad (2)$$

where $g(\cdot)$ is the tag generator function, k is the secret key and s is the modulated (reference) signal.

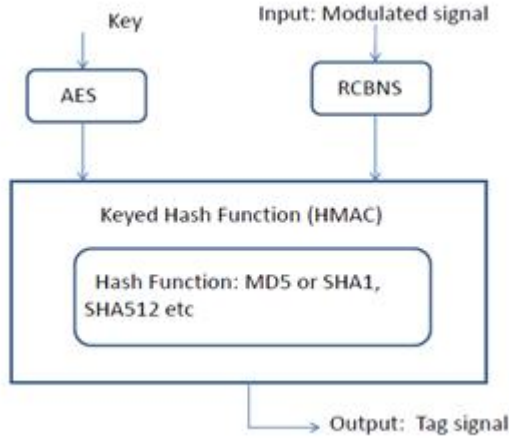


Fig. 3. Tag Generator

2.3 Superimpose

In this paper we are superimposing the modulated signal and tag signal as done in spread spectrum technique which spreads signal wave of message signal for same transmission power. In spread spectrum technique spreading code is pseudo random code which is combined with data signal via bit wise XOR operation in the same way the message signal is superimposed with tag signal in our case. Before superimposing, modulated signal is converted into binary form using RCBNS method, to obtain a tagged signal. The superimposed signal is given by:

$$x = \rho_s s + \rho_t t \quad 0 < \rho_s, \rho_t < 1 \quad (3)$$

As the message signal we assume the tags $E|t_k| = 0$ and $E|t_k|^2 = L$. ρ_s^2, ρ_t^2 are energy allocations of the message and tag signal. The tagged signal is in the binary form and hence converted into polar signaling. The baseband polar signaling is given by:

$$x(t) = \begin{cases} 1, & \text{for } t = 1 \\ -1, & t = 0 \end{cases} \quad (4)$$

This polar signal is an antipodal signal, since $x_1(t) = -x_2(t)$.

Before sending the tagged and reference signal through channel it is passed through the Root Raised Cosine (RRC) filter [5] which is frequently used as the transmitter and the receivers filter in case of digital communication. This is used for transforming base band signal into pulse wave form which is more appropriate for transmission through channel. This filter is used to reduce the ISI and also smoothing the sharp edges of the signal before transmitting through channel. The RRC filter is characterized by two values roll factor and reciprocal symbol rate.

Algorithm 1. Complex to Binary Conversion

```

1: BEGIN
2: CONVERT real and imaginary part into binary form.
3: if (Real and Imaginary)>0 then
4:   Place '0' in front of each 2-bit digit
5: else
6:    $q_i <-$ Get 3-bit for each part and combine them together starting with LSB as 3
       bit imaginary and put them together with 3 bits of real number and represent
       in a row.
7:   FIND equivalent values for every 'q' using table.
8:   //Conversion Table Generation
9:   MULTIPLY the positive sign to even rows and negative sign to the odd rows.
10:  COMBINE  $q_i$ 
11: end if
12: END

```

2.4 Channel

In this paper we have considered both wired and wireless channels for transmitting authentication signal. In wireless channels two types of channels are used they are, Rayleigh channel for outdoor communication and Rician channel for indoor communication.

Wired Channel. As depicted in Fig.4. A wired channel is usually modeled as AWGN (Additive White Gaussian Noise) channel. In this matched filter is used to recover the tagged signal and reference signal. The output of the matched filter is passed through decision device.

$$z = x_n + w_n \tag{5}$$

Wireless Channel. In case of wireless channel, the multipath is a propagation phenomenon that causes signals received at the receiver side in two or more paths. This affects the intersymbol interference to the signal. Rayleigh and Rician channel models are caused by multipath reception. In Rayleigh fading channel

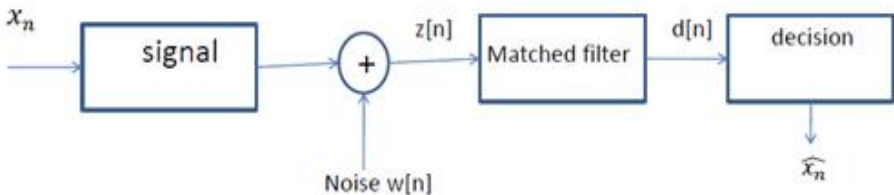


Fig. 4. Wired Channel Model

the receiver antenna receives large number of reflected and scattered waves. In the Rician channel the receiver antenna receives number of reflected and scatter waves and one line of sight wave.

$$z = h_n x_n + w_n \tag{6}$$

where h is the channel coefficient vector and w is the White Gaussian Noise and the average Signal-to-Noise Ratio (SNR) is $\gamma = \sigma_h^2 / \sigma_w^2$.

2.5 Signal Recovery

The Fig.5 shows on how the receiver should recover the signal. For recovering the signal we use the MMSE equalizer [5]. MMSE equalizer designs the filter to minimize $E[|e|^2]$ where 'e' is a error signal, which is the difference of filter output and transmitted signal. It is a linear equalizer and doesn't have any feedback. We can use the other equalizer, such as MLSE which is more complexity in design while we increasing quantization levels. The equalizer transfer function is given by

$$F(z) = \sum_{i=0}^{\infty} f(i)z^{-i} \tag{7}$$

$$d(n) = F(z)Z(n) = F(z)H(z)x_n + F(z)w(n) \tag{8}$$

Let $C(z) = F(z)H(z) = \sum_{i=0}^{\infty} c_i z^{-i}$. So the goal of the equalizer $F(z)$ is to clean up the ISI (Intersymbol interference) in $d(n)$ to achieve error free decision. Let $\hat{x}_n = \text{dec}(d(n)) = x_{n-u}$, where u is fixed delay in the equalizer output. $F(z)$ can be divided as

$$\begin{aligned} d[n] &= \sum_{i=0}^{\infty} c_i x_{n-i} + \sum_{i=0}^{\infty} f(i)w(n-i) \\ &= c_u x_{n-u} + \sum_{i=0, i \neq u}^{\infty} c_i x_{n-i} + \sum_{i=0}^{\infty} f(i)w(n-i) \end{aligned} \tag{9}$$

By considering the equalizer output, the overall distortion in $d(n)$ can be quantified as

$$d(n) - x_{n-u} = \sum_{i=0}^{\infty} c_i x_{n-i} + \sum_{i=0}^{\infty} f(i)w(n-i). \tag{10}$$

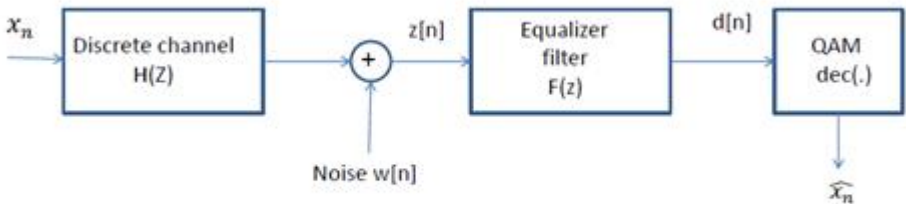


Fig. 5. Block Diagram of the Receiver

MMSE equalizer design minimizes the minimum mean square error $[d(n) - x_{n-u}]^2$. After getting $d(n)$ equalizer output it is passed through the decision device which gives the recovery signal \hat{x}_n . When the tagged and reference signal is recovered at receiver then using reference signal and tag generator we generate the tag signal and extract another tag signal from tagged signal as depicted in Fig.(2). Now both the tag signals are compared, if two tag signals differ, then the signal is not authentic and is rejected. If both are same, then the signal is considered to be authentic and the reference signal is demodulated.

3 Simulation Results and Discussion

The entire proposed solution has been simulated in Matlab7.6. Here we have taken a randomly generated input which is to be modulated using 16-QAM technique. This modulated signal is passed through tag generator function, which generates the tag signal. In tag generator, SHA512 is used to generate an intermediate tag, which is in turn provided to HMAC along with shared secret key to obtain the final tag signal. This tag signal is superimposed with the modulated signal which is called tagged signal. The tagged and reference signal is passed through the Root-Raised cosine filter with a roll-off factor of 0.5. Before transmitting these two signals through channel it is passed through low pass filter which is a Chebyshev Type2 filter with parameter of order 12, stopband gap is 20 dB and bandwidth is 0.75/THz, then it is sent through the channels. For our simulation we have considered four test cases:

- (i) Tagged signal and reference signal is transmitted to aware receiver Bob without any errors. In Fig.6 (a) shows that Bob receives the same signal which is sent by Alice. Fig.6 (b) depicts a tagged signal as received by Bob is same as sent by Alice. In Fig.6 (c) depicts the comparison of two tag signals that is generated by Bob using reference signal and tagged signal which are same. Hence considering the received signal to be authentic.
- (ii) Transmitter transmits tagged signal and reference signal and Eve intercepts the signal with wrong secret key. Fig.7 shows that, Eve intercepts the signal which is sent by Alice to Bob. Assuming Eve does not know the mechanism and shared secret key that is used for communication. In Fig.7 (a) Eve intercepts the reference signal that is being transmitted by Alice. Fig.7 (b) shows the tagged signal that is been intercepted by Eve. In Fig.7(c) shows the two compared tag signals, which are not same, hence the received signal is not authentic and therefore gets rejected.
- (iii) Transmitter transmits tagged signal and reference signal and Eve intercepts the signal and changes the signal and forwards it to the Bob. Bob rejects the signal because it is not an authentic signal. Fig.8 shows that, Eve intercepts the signal which is sent by Alice to Bob. Assuming Eve knows the mechanism to generate the tag signal and doesn't know the shared secret key. Here Eve changes the reference signal and sends it to Bob. In Fig. 8(a) depicts the comparison of the actual message signal which is sent by Alice and reference signal received by Bob are not same. Fig.8 (b) depicts

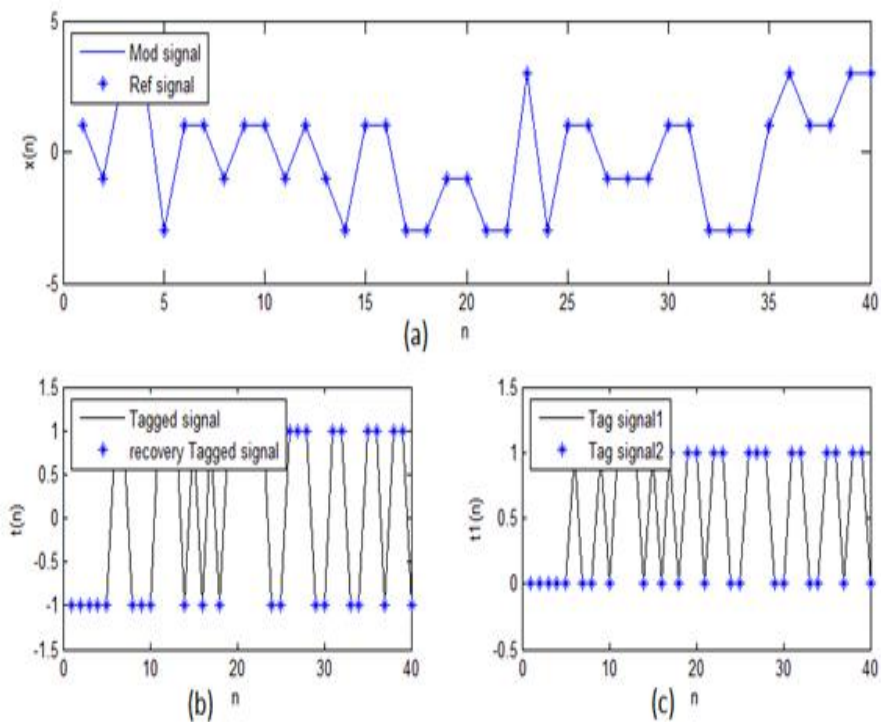


Fig. 6. Actual Signal Received by Bob from Alice

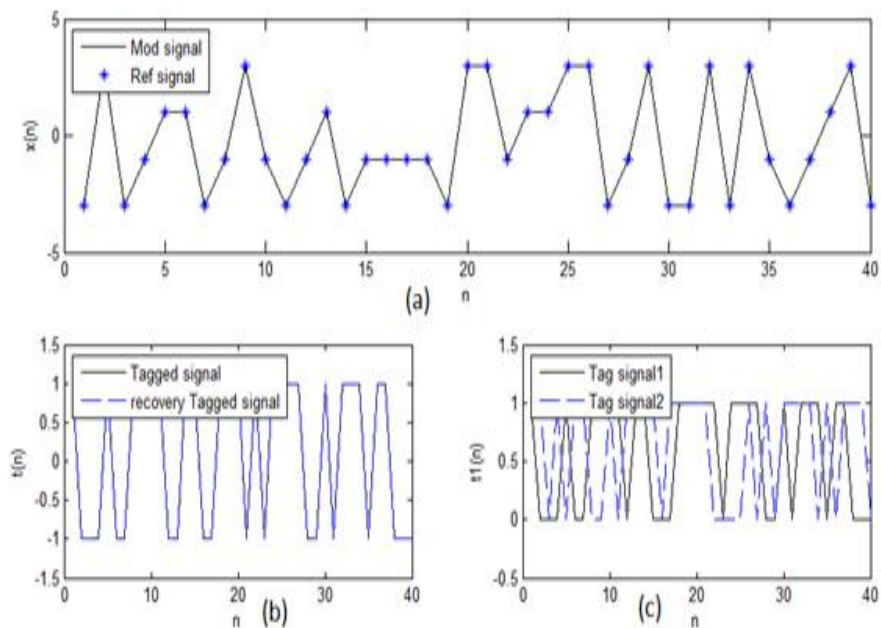


Fig. 7. Eve Intercepts the Signal Transmitted by Alice

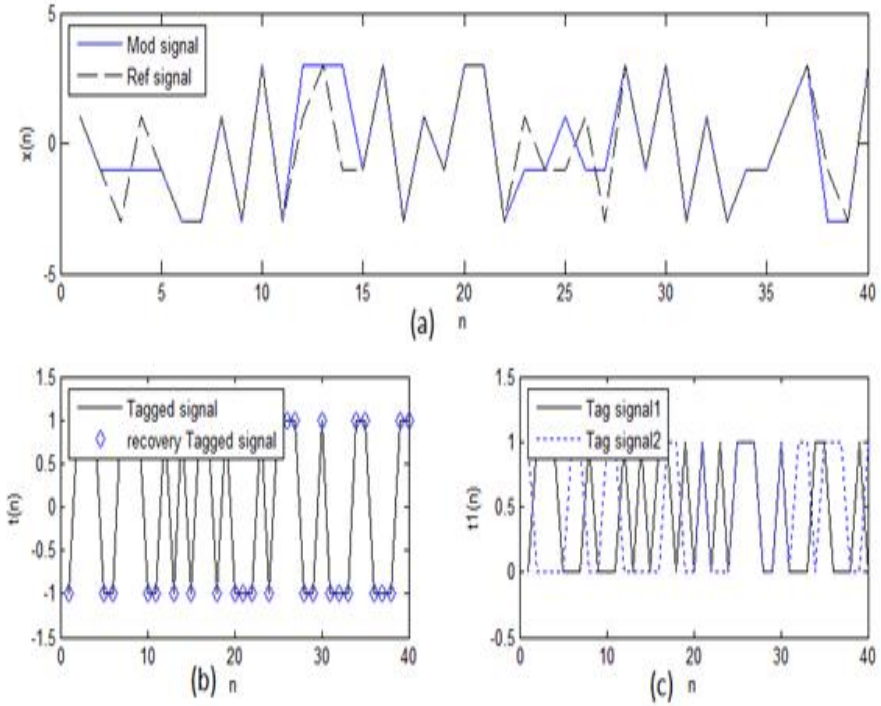


Fig. 8. Signals Received by Bob from Eve

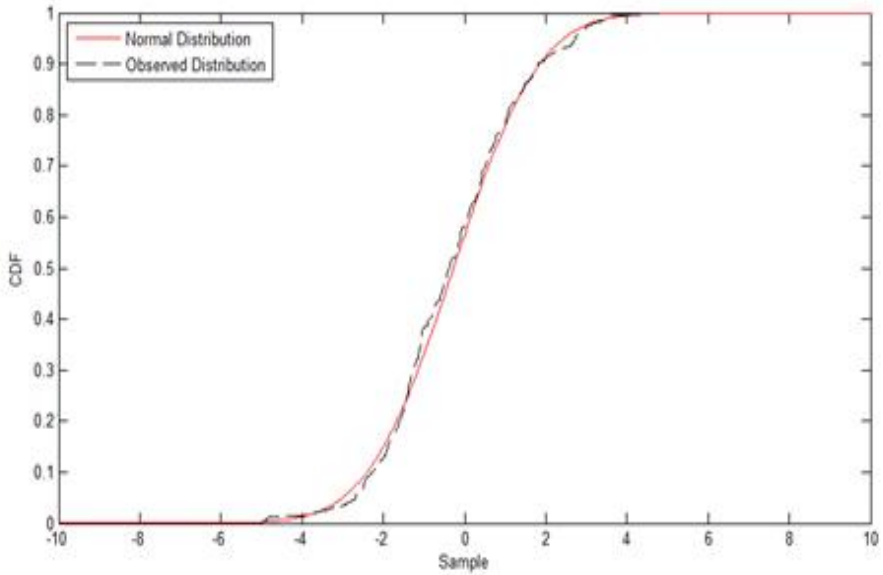


Fig. 9. Noise CDF, TNR = 0dB, 2-bit tag

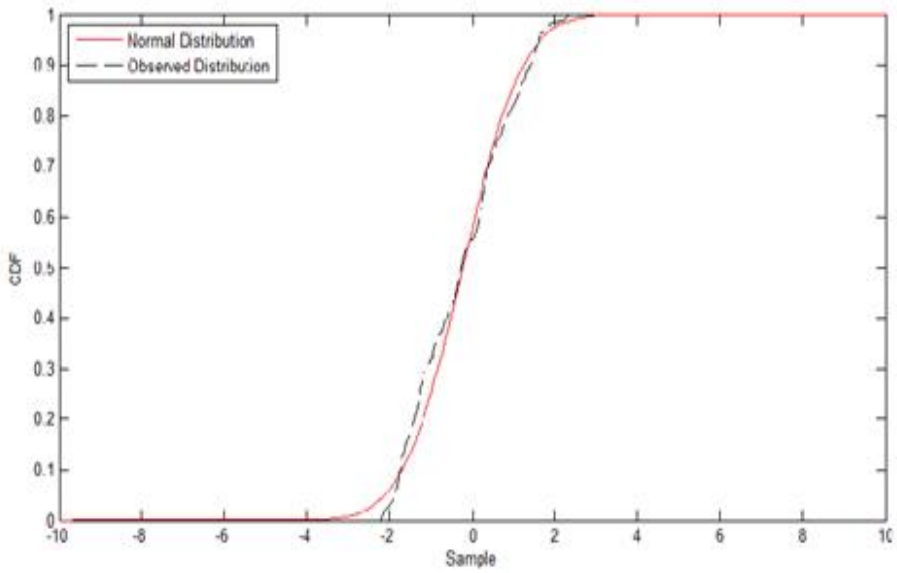


Fig. 10. Noise CDF, TNR =10dB,2-bit tag

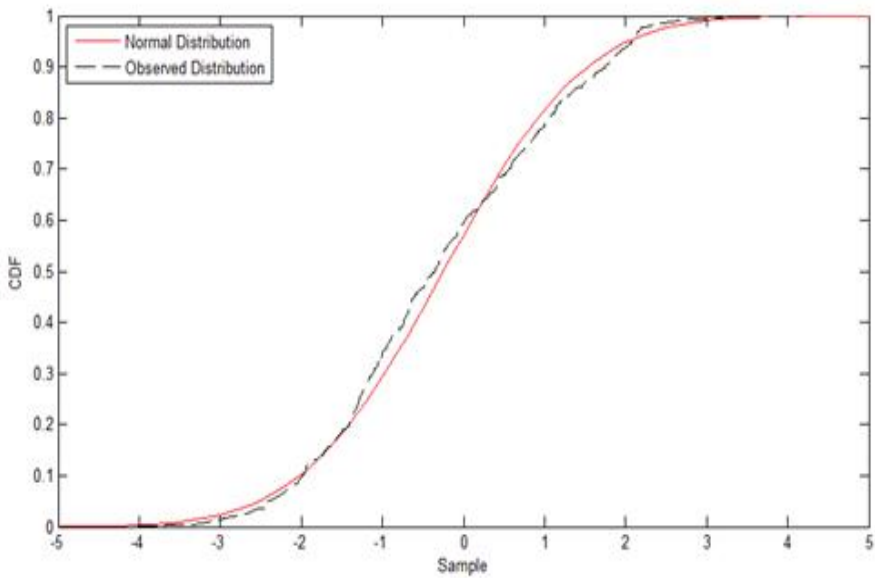


Fig. 11. Noise CDF, TNR =10dB, 1-bit tag

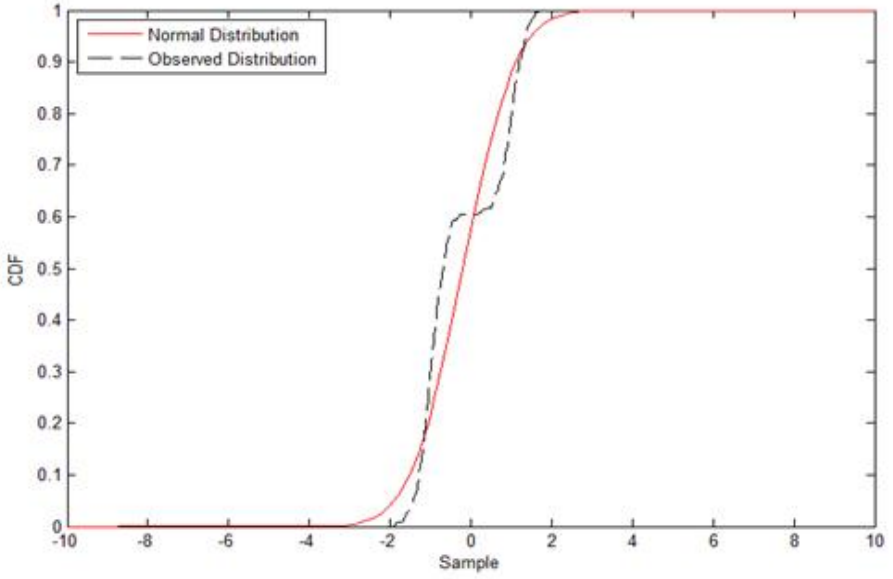


Fig. 12. Noise CDF, TNR = 0dB, 1-bit tag

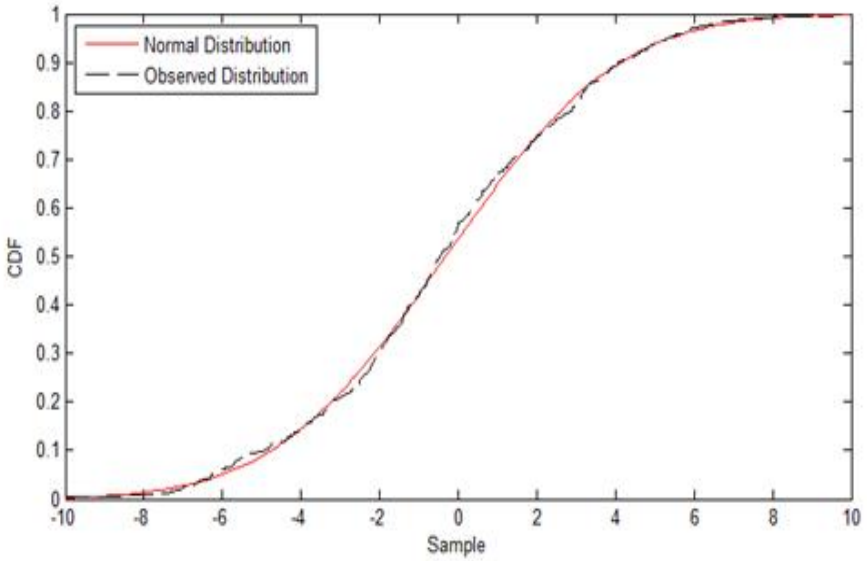


Fig. 13. Noise CDF, TNR = -10dB, 1-bit tag

that the tagged signal sent by Alice and recovered at Bob are same. By using recovery tagged signal and reference signal, Bob generates two tag signals. Fig.8(c) shows the comparisons of the two signals, as the signals are not same then Bob knows that the message signal is not authentic.

- (iv) The tag signal should be covert that is, it should not be easily detectable. For this we are applying the goodness-fit-test to the tag signal. Here goodness-fit-test used is Lilliefors [11] tests which provides a well-known class of anomaly detection algorithms. Considering a simple experiment, we ignore the channel effects and tag is observed in the AWGN $y_k = t(k) + w_k$. Let the TNR (Tag noise ratio) be σ_t^2/σ_w^2 where $\sigma_t^2 = E[|t_k|^2]$. This goodness-fit-test compares the normal cumulative distribution function (CDF) with estimated mean and variance from observations and empirical CDF. Lilliefors test does not detect the anomalous signal when we increase TNR value. This means it distinguishes with normal and observed distribution with estimated mean and variance.

a Noise CDF, with TNR = 0dB, 10dB, 2-bit tag taking one of the values -1.51,-0.453, 0.453, 1.51.

Fig.9 and 10 shows the empirical versus normal CDFs when the 512 2-bit i.i.d tag symbols are drawn and observed with TNR = 0dB, 10dB. Lilliefors test is unable to detect distinguish between the CDFs in Fig.9 but in Fig.10 there is difference between the empirical versus normal CDFs.

b Noise CDF, with TNR = 0dB, 10dB and -10dB, 1-bit tag.

Now, suppose that each tag symbol is represented by one of two equiprobable and polar values $\pm\sigma_t$. Fig.11, 12 and 13 shows the empirical versus normal CDFs when the tag has 1-bit symbol and varied TNR values. At TNR 10dB and 0dB in Fig.11 and 12 the observed CDFs becomes distinguishable with the normal CDFs. When we lower the TNR to -10 dB in Fig.13, the observed distribution becomes indistinguishable from normal distribution. By the above figures we can improve the covertness by transmitting the tag signal at lower TNR or low power or tag can follow the noise-like distribution.

4 Conclusion

Authentication is a security service that checks the identity of the party at the other end of the line. We have generated tag signal and transmitted the tag signal concurrently with data without using any extra bandwidth and transmission cost on wired and wireless channels. Here we have shown how the generated tag signal could be superimposed with data signal. We depict simulation results for the above stated four test case scenarios. The obtained results showed that physical layer authentication is stealthy, secure and could be effectively implement in real time communication. In future we would like to implement the propose schema in different wireless channels like Nakagami fading channel, Gauss-Markov channel and observe the behavior of our solution proposed.

References

1. Simmons, G.: A survey of information authentication. *Proceedings of the IEEE* 76(5), 603–620 (1988)
2. Menezes, A.J., van Oorschot, P.C., Vanstone, S.A.: *Handbook of Applied Cryptography*, 1st edn. CRC Press, New Delhi (2001)
3. Cover, T.M.: Broadcast channels. *IEEE Trans. Info. Theory* IT-18(1), 2–14 (1972)
4. Cox, I.J., Miller, M.L., McKellips, A.L.: Watermarking as communications with side information. *Proceedings of the IEEE* 87(7), 1127–1141 (1999)
5. Proakis, J.G.: *Digital Communications*, 5th edn. MacGraw-Hill, New Delhi (2001)
6. Supangkat, S.H., Eric, T., Pamuji, A.S.: A public key signature for authentication in telephone. In: *2002 Asia-Pacific Conference on Circuits and Systems, APCCAS 2002*, vol. 2, pp. 495–498 (2002)
7. Kleider, J., Gifford, S., Chuprun, S., Fette, B.: Radio frequency watermarking for ofdm wireless networks. In: *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2004)*, vol. 5, pp. 397–400 (May 2004)
8. Yu, P., Baras, J., Sadler, B.: Physical-layer authentication. *IEEE Transactions on Information Forensics and Security* 3(1), 38–51 (2008)
9. Zaini, H., Deshmukh, R.: Complex number representation in rcbns form for arithmetic operations and conversion of the result into standard binary form. In: *Proceedings of SoutheastCon.*, pp. 111–116. IEEE, Los Alamitos (2003)
10. Behrouz, A.F.: *Cryptography and Network Security*. Tata MacGraw-Hill, New Delhi (2007)
11. Salkind, N.: *Encyclopedia of Measurement and Statistics*, 1st edn. Sage Publication, Thousand Oaks (2007)

Cloud SaaS: Models and Transformation

Ritu Sharma and Manu Sood

Himachal Pradesh University Summer Hill, Shimla 5,
Himachal Pradesh, India
rituchetan@gmail.com, soodm_67@yahoo.com

Abstract. Cloud computing is a new and upcoming computing model where the computational resources are offered as services, remotely over a network. The software applications running as cloud SaaS are targeted on different platforms. The constantly evolving software technologies may result in the obsolescence of the legacy applications. It, therefore, becomes an essential requisite to adopt a software development methodology that could alleviate the undesirable effects of technology change. The Model Driven Architecture (MDA) approach becomes the obvious choice where the models drive the software development process. These models are defined at different levels of abstraction; and automated tools are used for model-to-model and model-to-code transformations between the levels. In this paper, the authors illustrate an MDA-based approach for developing cloud software applications making them more robust, flexible and agile, in the wake of changing technologies.

Keywords: Cloud computing, Cloud SaaS, Model-Driven Architecture (MDA), Computation Independent Model (CIM), Platform Independent Model (PIM), Platform Specific Model (PSM), Model Transformation.

1 Introduction

Cloud computing, a recent computing paradigm in the Information and Communication Technology (ICT) field, offers computational resources such as – storage units, processing units, computing platforms, development environments and other infrastructure – as services over an intranet or Internet, on subscription or metered usage basis. These resources in the cloud are dynamically scalable and often virtualized. The services can be accessed on demand using a browser running on a thin client or even a mobile phone.

The rapid advancements in the field of ICT in recent years entail the development of software solutions in a manner that is independent of the technology change. Model Driven Architecture (MDA) is a software development approach where formal models defined at different levels of abstraction, are used as the prime artifacts throughout the development process. The transformation of models from one level of abstraction to another, or the transformation of models to executable code is performed using (semi)automated transformation tools.

This paper attempts to incorporate the model-driven software development methodology in the development of cloud SaaS (Cloud Software-as-a-Service).

Section 2 discusses the concept of cloud computing. Section 3 briefly discusses the basics of MDA. Section 4 illustrates the Computation Independent Model (CIM), Platform Independent Model (PIM), and Platform Specific Model (PSM) of a cloud software application taken as example, and also discusses the transformation definition for its PIM to Relational PSM transformation. Section 5 draws the conclusion of the paper and the future work undertaken by the authors.

2 Cloud Computing

The term 'cloud' is a metaphor for the Internet which is graphically represented as a cloud. Thus a cloud, in the given context, refers to a complex, internet-based infrastructure of hardware and software components. Cloud computing encapsulates a variety of services such as Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS), Infrastructure-as-a Service (IaaS), and Hardware-as-a-Service (HaaS) that can be deployed in a private, public or hybrid cloud [1, 2].

A *cloud SaaS* refers to a multi-tenant platform that uses common resources and a single instance of both – the object code of an application as well as the underlying database – to support multiple customers simultaneously. It is based on the Application Service Provider (ASP) model and heralds a new wave in application software distribution. Some of the key providers of this service are SalesForce.com, NetSuite, Oracle, IBM, Microsoft etc [3, 4].

A variety of technologies and standards that led to the emergence of cloud computing include distributed computing, grid computing, virtualization, Web services, Web Service Definition Language (WSDL), Simple Object Access Protocol (SOAP), Universal Description Discovery and Integration (UDDI), Service Oriented Architecture (SOA), Software-as-a-Service (SaaS) etc.

Cloud computing has changed the outlook of the enterprises especially the small and medium enterprises, towards their business solutions. They can now utilize the cloud services without incurring any upfront infrastructure costs. As per demand, the location-independent resources in the cloud are acquired and released with great elasticity to meet the requirements of the service consumers. The resource usage is measured and the consumer pays for the services rendered to him.

3 The Model-Driven Architecture (MDA) Approach

Model Driven Architecture (MDA) is a standardization effort by OMG which aims at reducing the gap between business modeling and software development significantly, by ensuring that business models drive the application development. While traditional software design and development processes create applications for deployment to a specific technology platform, MDA introduces higher levels of abstraction, enabling organizations to create *models* that are independent of any particular technology platform. The entire software development process in MDA is model-driven with models as the primary artifacts for understanding, design, construction, deployment, operation, maintenance and modification of a system [5]. Although the high level frameworks such as J2EE, .NET and other Web services also raise the level of abstraction, but they still focus on computer-level concepts rather than business level concepts.

The models in MDA are defined at three levels of abstraction – the Computation Independent Model (CIM), the Platform Independent Model (PIM) and the Platform Specific Model (PSM). A CIM is a software independent model and specifies the business logic of the application. A PIM describes a software system that supports the business independent of any implementation technology. A PSM specifies the system in terms of implementation constructs that are specific to the implementation technology. A single PIM is transformed into one or more PSMs, each PSM being specific to the technology platform on which the system would finally be implemented.

The success of MDA approach resides in automating the model-to-model and model-to-code transformations. The transformation tools automatically transform higher-level, platform-independent business models into lower-level platform-specific models and finally into executable code. A transformation tool contains the transformation definition. A transformation definition is a set of transformation rules that together describe how a model in the source language can be transformed into a model in the target language [6]. Figure 1 depicts a PIM to PSM transformation.

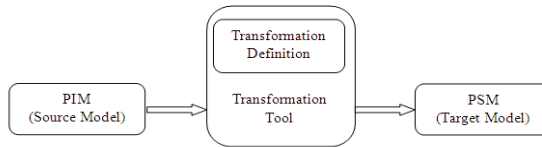


Fig. 1. PIM to PSM Transformation

4 A Cloud SaaS: CIM, PIM and PSM

US NIST (National Institute of Standards and Technology) [7] defines Cloud SaaS as a service model in which the applications running on a cloud infrastructure and managed by the service provider, are offered as services to the consumers on a pay-per-use basis. These services rendered by the provider are accessible from various client devices through a thin client interface such as a web browser, and may require limited user-specific configuration settings. The application may be simple one performing a single function or a complex one performing a set of related functions.

With constantly evolving technologies, it becomes inevitable to adopt a software development methodology that would not only decouple the cloud SaaS from the undesirable effects of technology change, but also enhance their longevity. An MDA-based development of cloud SaaS (application) will enable defining these services in a technology-independent manner and will play a significant role in improving the quality of cloud software services, making them more robust, flexible and agile [8]. Encapsulating business logic in a manner that is independent of the technical mechanisms will formally capture the essence of the applications; and will also make it possible to reuse them in a variety of contexts [9, 10].

We illustrate the model-driven approach to develop a cloud SaaS (application), with the help of an example – the Online Hotel Reservation System (OHRS). The OHRS, a software application running as a service in the cloud, may be utilized by

any small or medium scale enterprise (SME), on a payment basis. The proposal would also be suitable for the entrepreneurs who wish to start up with a hotel business on a small scale and do not wish to initially invest huge capital in purchasing, installing or developing the hardware and software for the purpose. The OHRS would perform a variety of tasks for its users (both customers and hotel administration) which include determining the availability of rooms in the hotel, online reservations, online cancellations, generating arrival chart, generating reports for decision-making, modifying room tariffs etc.

Although, the MDA approach does not restrict itself to the use of UML (Unified Modeling Language) for modeling the software applications, the authors are using UML to model different aspects of the system.

4.1 Computation Independent Model (CIM) for OHRS

A CIM specifies the computation independent view of the system. It captures the requirements of the system in a vocabulary familiar to the domain practitioners. The loosely defined requirements of the business process are distilled into requirements model without losing important details. As part of the CIM, the Use Case diagram and the Activity diagram available in UML have been used to model the functional requirements of the system under consideration.

Each *use case* in the use case diagram for OHRS, depicted in Figure 2, captures a piece of functionality that the system provides. It specifies ‘*what*’ a system is supposed to do. The characteristics of the actors interacting with the OHRS are:

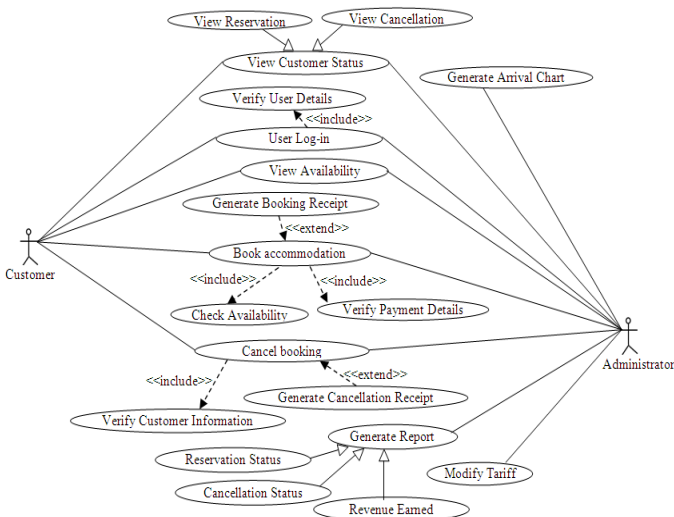


Fig. 2. Use Case Diagram for OHRS

- A Customer is a person who accesses the OHRS to view the availability status or to book/cancel an accommodation in the hotel.
- An Administrator is the hotel personnel who accesses the OHRS for not only viewing the availability status or booking/cancelling accommodations for its customers over the counter, but also to generate various reports for decision making or for modifying the tariffs of unit types.

The Activity diagram specifies how a system would accomplish its goals. It models a set of coordinated tasks that achieve a business goal. The high-level actions chained together to represent the business process of the OHRS is depicted in the Activity Diagram in Figure 3.

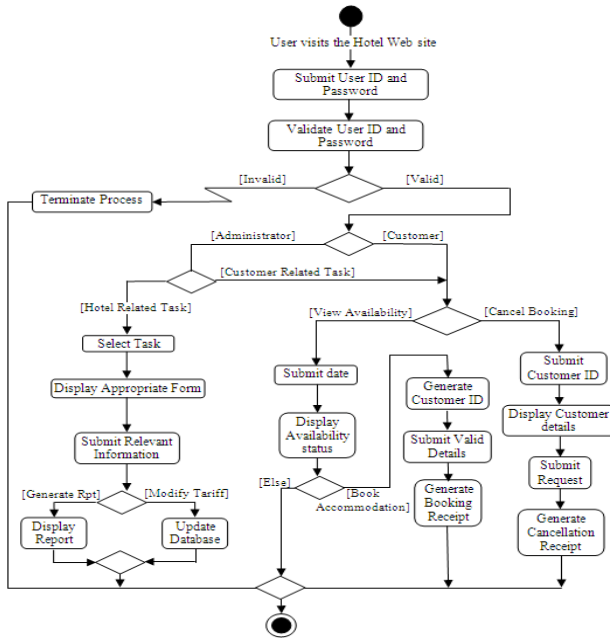


Fig. 3. Activity Diagram for OHRS

The various steps in the process may be listed as:

1. A user (customer or administrator) logs into the hotel website with a valid user ID and password.
2. A customer may perform various tasks such as viewing the availability status, book accommodation or cancel a previously booked accommodation.
3. In order to view the availability status, the customer submits the date in response to which the system displays the availability status of the hotel for the next fifteen days.
4. The customer may then either quit the application or may proceed to book an accommodation based on its availability.

5. In response to the book accommodation action, the system generates a unique ID for the customer.
6. Next, the customer submits his details such as name, address, phone, email address, unit type, number of units, book-from date and book-to date and payment details.
7. A booking receipt bearing all the details is generated for the customer and the process terminates.
8. A customer may also cancel a previous booking, in which case he submits the customer ID that was generated at the time of booking.
9. In response to step 8, the system displays the details of the customer.
10. The customer submits the cancellation request.
11. A cancellation receipt is generated for the customer and the process terminates.
12. An administrator, after successful log-in, is able to perform all the tasks that a customer does and can also generate reports or modify tariffs by submitting the required information.

In an ideal MDA approach the requirements model should be simply submitted to the generators that would produce the required systems. But, in practice the requirements model need to refined further into a computational model that a generator can process.

4.2 Platform Independent Model (PIM) for OHRS

A PIM specifies the system at a higher level of abstraction as compared to a PSM. There is only one single PIM for the software system that can be mapped to several different PSMs targeted on different platforms by defining platform-specific transformation rules for the purpose.

The PIM for the OHRS is depicted in Figure 4, by means of a class diagram. The classes, their attributes and operations, and the associations among various classes are shown in the model. The get and set operations have been intentionally excluded from the diagram in order to keep it simple. A multiplicity adorns each association. This PIM defines the static aspects of the OHRS application through a static view. Though this model reflects the technicalities of the system, it is non-committal to the platform that would implement and host the system.

The various classes in the class diagram for specifying the PIM for OHRS are – Hotel, Customer, UnitType, Administrator, User, Booking and Payment. Booking and Payment are association classes. The associations among the classes are defined as under:

- A hotel is a composite aggregation of unit types.
- A hotel has one or more unit types.
- A hotel may have zero or more customers.
- A customer with a unique ID belongs to only one hotel.
- A hotel has one or more Administrators.
- An administrator belongs to only one hotel.
- A customer may book one or more unit types.
- A unit type may be booked by zero or more customers.
- An administrator may book zero or more unit types.

- A unit type may be booked by zero or more Administrators.
- A customer makes one or more payments to the hotel.
- A payment is related to only one customer.
- A unit type is related to zero or more bookings.
- A booking is related to only one unit type.
- A user may be a customer or an administrator.
- A customer is essentially a user of the system.
- An administrator is essentially a user of the system.

The PIM of the cloud SaaS describes the attributes and operations in a manner that is entirely independent of any programming language or operating system.

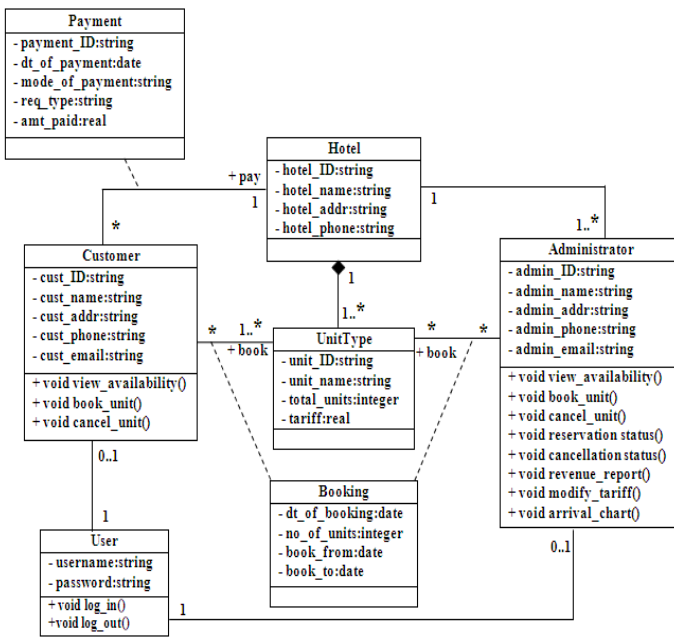


Fig. 4. Platform Independent Model for OHRS

4.3 Platform Specific Model (PSM) for OHRS

A PSM describes the technology specific details for the target platform. One or more PSMs can be derived from a single PIM using automated transformation tools.

The PIM of OHRS is transformed into a Relational PSM based on certain transformation rules defined in section 4.4. The Relational PSM specifies the database and is described by a relational model depicted in an Entity-Relationship diagram in Figure 5.

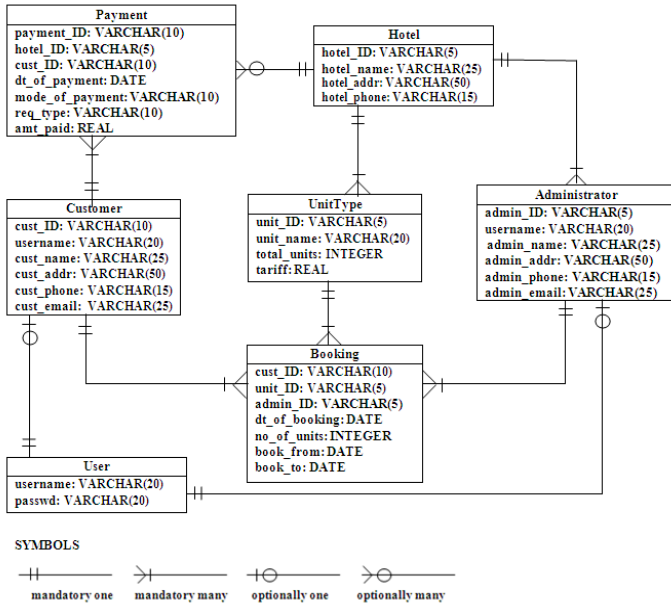


Fig. 5. Relational Model for OHRS (PSM)

In [8], the authors discuss a model-driven approach to obtain a Java PSM from a PIM of a cloud SaaS taken as example.

4.4 Transformation Definition for OHRS: PIM to Relational PSM

A transformation definition is comprised of transformation rules that describe how one or more constructs in the source language can be transformed into one or more constructs in the target language. The transformation rules defined here take care of a consistent object-relational mapping [6]. These rules describe how the elements in PIM can be mapped to elements in Relational PSM. Based on the various elements in the PIM, the transformation rules are categorized into:

A. Transformation Rules for Data types: A basic data type is mapped to corresponding data type in the relational model according to the following rules:

- i. A UML string is mapped onto a SQL VARCHAR(n).
- ii. A UML integer is mapped onto a SQL INTEGER.
- iii. A UML real is mapped onto a SQL REAL.
- iv. A UML date is mapped onto a SQL DATE.

A data type may be a simple data type or a derived data type.

a. Transforming Simple Data types: A simple data type is mapped directly to a column in a table.

b. Transforming Derived Data types:

- i. A UML data type that has no operations, such as struct and array, is mapped onto a number of columns, each representing a field in the derived data type.

- ii. A UML data type that is a class is mapped to a table itself. The column holds a reference (foreign key) to a key value in the other table.
- B. Transformation Rules for Classes and Attributes:
- a. Each UML class is transformed into a Relation/Table with the same name.
 - b. Each UML attribute is transformed into a field (column) in the table with the same name.
 - c. When the type of attribute is not a data type but a class, the field in the table holds a foreign key to the table representing that class.
- C. Transformation Rules for Associations :
- a. Associations in the UML model are transformed into a foreign key relation in a database model, possibly introducing a new class.
 - b. An association class represents a relationship between two classes. It is transformed into a table with foreign keys in this new table referring to the two related tables.
 - c. The multiplicities of an association from class A to class B may be:
 - The multiplicity at A is zero-or-one,
 - The multiplicity at A is one, or
 - The multiplicity at A is more than one.

The same holds for the multiplicity at B resulting in nine different combinations of multiplicities at both the ends. A pseudocode for the rule may be expressed as:

```

if the association A to B is adorned by an
  association class or the multiplicity at
  both ends is more-than-one
then create a table representing the
  association class or the association and
  create foreign keys in this new table
  referring to the related tables
else if the multiplicity at one end is
  zero-or-one
  then create a foreign key in the table
  representing the class at that end,
  referencing the other end
  else /* the multiplicity of the
  association is one-to-one */
  create a foreign key in one of the
  tables, referencing the other end
endif
endif

```

5 Conclusion and Future Work

Cloud computing is fast emerging as a computing paradigm where the computations would be performed using third-party hardware and software resources. In the wake of constantly changing technologies, the software applications in the cloud, if developed for specific platforms, would soon become obsolete. The authors in this paper therefore stress on the need to leverage the MDA approach in the development of cloud SaaS.

Incorporating MDA approach in developing a cloud SaaS has two-fold advantage – firstly, the need to redefine the structure, behavior and functionality of the application for different technology-specific platforms is ruled out as the same PIM can be used to generate different PSMs using transformation tools; and secondly, as the technologies evolve, the same PIM can be used to develop PSMs and code specific to new or evolving technologies. This is expected to minimize the time, cost and efforts in developing cloud SaaS and enhance the Return on Investment. Thus, the developers as well as the service consumers will be able to reap the benefits of both – cloud computing and MDA.

The authors have recently developed a transformation tool based on the transformation definition discussed in this paper. Enhancing the research further, the authors are presently working on a model-driven approach to ensure interoperability among software services identified in a cloud, in the light of SOA.

References

- [1] Youseff, L., Butrico, M., Da Silva, D.: Toward a Unified Ontology of Cloud Computing. In: Grid Computing Environments Workshop, pp. 1–10 (2008) ISBN: 978-1-4244-2860-1, doi: 10.1109/GCE.2008.4738443
- [2] Maggiani, R.: Cloud computing is changing how we communicate. In: IEEE International Professional Communication Conference, pp. 1–4 (2009)
- [3] Rimal, B.P., Choi, E., Lumb, I.: A Taxonomy and Survey of Cloud Computing systems. In: Fifth International Joint Conference on INC, IMS and IDC, pp. 44–51 (2009)
- [4] Foster, I., Zhao, Y., Raicu, I., Lu, S.: Cloud Computing and Grid Computing 360-Degree Compared. In: IEEE Grid Computing Environments Workshop, pp. 1–10 (2008)
- [5] Miller, J., Mukerji, J.: MDA Guide Version 1.0.1, <http://www.omg.org/docs/omg/03-06-01.pdf>
- [6] Kleppe, A., Warmer, J., Bast, W.: MDA Explained: The Model Driven Architecture: Practice and Promise. Pearson Education, Inc., London (2003)
- [7] Mell, P., Grance, T.: The NIST Definition of Cloud Computing. Version 15 (July 10, 2009), <http://thecloudtutorial.com/nistcloudcomputingdefinition.html>
- [8] Sharma, R., Sood, M.: Modeling Cloud Software-as-a-Service: A Perspective. In: International Conference on Network Communication and Computer (ICNCC 2011), pp. 170–174 (2011) ISBN: 978-1-4244-9550-4
- [9] Frankel, D.S.: Model Driven Architecture: Applying MDA to Enterprise Computing. Wiley Publishing Inc., Chichester (2003)
- [10] Frankel, D., Parodi, J.: Using MDA to develop Web Services, 2nd edn. IONA Technologies PLC (April 2002)

Future Robotics Memory Management

S. Vijaykumar¹ and S.G. Saravanakumar²

¹ Thiagarajar School of management
Thirupparankundram
Madurai 625 005, Tamil Nadu, India
indianid@gmail.com,
Project6thsense@googlegroups.com
<http://www.indianid.6thsense.us>
² Sastra University, Kumbakonam-612001
Tamil Nadu, India
saravanakumarsg@gmail.com

Abstract. This paper deals the brand new future memory management in robotics field. It is a massive challenging issue of this proposal because till now existing database system did not support robotics. While designing the database it is very complex to make a robotic database made with complex designs for the effective data retrieval so here we provide a new concept called NOSQL. It comes under the non-relational database type it can handle the large volumes of data and it can easily process it. It does not have rigid schema design we can apply multiple operations on the single row and in addition we included map reduce concepts it can splits the job to the respective workers so that it can map the data easily in such way that the robots can think, through this map reduce approach we obtain parallelism to processing the large dataset.

Keywords: NOSQL, DBMS, RDBMS, Unstructured Database, Google, Big table, Humanoid, Map reduce.

1 Introduction

In the modern age robotics, we are facing very big problem to manage and retrieve information. When the embedded system we can only able to perform some task only like playing football robots, manufacturing robot, etc.in embedded system we are using multiple sensor to do various task and from sensor we give specific task for robots. But when the matter comes to making a robot like human beings humanoids we have to consider n number of things like intelligence up to instructions. Consider an middle finger have tree joints for that we want to give tree part of instruction and each part has ten states fuzzy sets (Oto1) for the movement and these have property like movement fast this way for each items we have to manipulate billions of instruction instead of that we go with an sensor on that case also we have an minimize the instruction regarding to that we want alternate solution to done an massive process and with massive intelligence. How it is possible? Because for that we need huge amount of memory, processing capacity, huge instruction, upgradable capability,

storage memory and a final important thing is data management. These are the key blocking things in the humanoid making process. But today we have an ability to achieve this by using a techniques and technology like NOSQL, fuzzy logic, Map Reduce, etc. Here we take a challenge to make a humanoid with amazing intelligence with massive processing technique. From this method we can able to store and retrieve yopta byte (1024) of information. Using this NOSQL we can able to retrieve that information in a quicker and efficient manner.

1.1 Database

A database is any collection of related data. And the restrictive of a database is a persistent, logically coherent collection of inherently meaningful data, relevant to some aspects of the real world [26].

1.2 Database Management System

A database management system (DBMS) is a collection of programs that enables users to create and maintain a database. According to the ANSI/SPARC DBMS Report (1977) [26].

1.3 Relational Database

A database that treats all of its data as a collection of relations and the characteristics of relations are [26].

A kind of set, a subset of a Cartesian product and an unordered set of ordered tuples

1.4 Problem with RDBS

The important problem with a RDBMS is difficult to scale bulk amount of data. they have facing 3 TB for "Green Badges", on that way Facebook handles 100 TB for inbox search and EBay handles 2PB and twitter handles 2PB every day for user images so the relational base are difficult to handle this much amount of data due to rigid schema design is the cause for this failure and we know server crash also happen due to data management sometimes the geo informatics service server also crash because of DBMS failure think it is the small amount of information when we compare with an humanoid knowledge information because maps are part of an humanoid information because the situation may occur to store more than 500TB to store human face images but in RDBMS how do we able update human face changes it is also an major restriction .

2 Definition for NOSQL

Next Generation Databases address some of the following being non-relational, distributed, open-source and horizontal scalable more nodes can be added. The original intention has been modern web-scale databases. NOSQL was first developed in the late 1990's by CarloStrozzi. The movement began early 2009 after it's growing rapidly.

2.1 Translation Table

This translation table explains you to know the NOSQL properties by its equivalent older meaning

Table 1. Keyword Translation

OLD NAME	NEW NAME
Hash file	Key-Value Store
Hierarchical file (HSAM, HDAM)	BigTable
Parent node	Column family
Local autonomy	Partition tolerant
Horizontal partition	Sharding
non-ACID (atomic, consistent, isolated, durable)	BASE (basically available, soft state, eventually consistent)

2.2 Characteristics

NOSQL normally doesn't have an ACID property like (atomicity, consistency, isolation, durability), no join operation, special of the NOSQL is schema-free, replication support, easy API, eventually consistency, and more. So the misleading term "NOSQL" (the community now translates it mostly with "Not Only SQL"). And it is structured storage and usually has a collection of tables with structured data (most probably like a hash table or a dictionary) then no need to map object-oriented designs into a relational model.

Examples Google's BigTable, Amazon's Dynamo. Cassandra (used in Facebook's inbox search) and HBase (Apache) are open source

2.3 CAP Theorem and NOSQL

[28]CAP (FOR NOSQL DATABASES)(FOR EASY SCALABILITY)

- **CONSISTENCY:** All database clients see the same data, even with concurrent updates.
- **AVAILABILITY:** All database clients are able to access same version of the data and easy scalability
- **PARTITION TOLERANCE:** The database can be split over multiple servers.

2.4 CORE NOSQL SYSTEMS

NOSQLS Systems where many in types but these where the core types of NOSQL Systems

1. Store / Column Families
2. Document Store
3. Key Value / Tuple Store
4. Eventually Consistent Key Value Store
5. Graph Databases

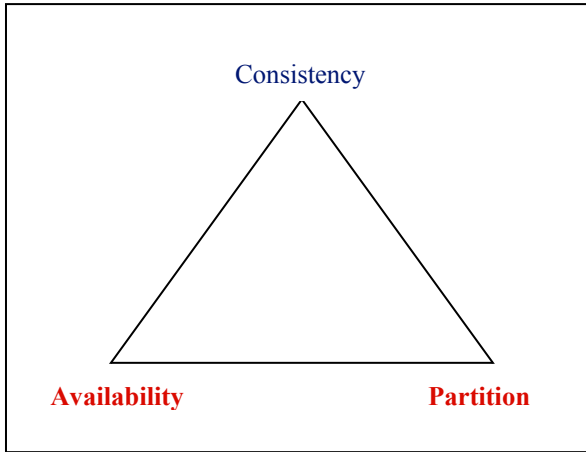


Fig. 1. CAP Theorem Satisfaction

Table 1. Data and Query Model Table

	Data Model	Query API
Cassandra	Column family	Thrift
CouchDB	Document	Map/reduce views
HBase	Column family	Thrift,REST
MongoDB	Document	Cursor
Neo4J	Graph	Graph
Redis	Collection	Collection
Riak	Document	Nested hashes
Scalaris	Key/Value	Get/put
Tokyo Cabinet	Key/Value	Get/put
Voldemort	Key/Value	Get/put

2.5 Data and Query Model

There is a lot of variety in the data models and query APIs in NOSQL databases.

2.6 Persistence Design

This are many number of persistence design avail today but above I give some famous model. It gives you a various choice to implement NOSQL depend on your need here below I take one GRAPH model Neo4J to give an view about that from that we can able to analyses the scenario and implement on it.

Table 2. Data Storage Design Table

	Persistence Design
Cassandra	Memtable/SSTable
CouchDB	Append-only B-tree
HBase	Memtable/SSTable on HDFS
MongoDB	B-tree
Neo4J	On-disk linked lists
Redis	In-memory with background snapshots
Scalaris	In-memory only
Tokyo Cabinet	Hash or B-tree
Voldemort	Pluggable(primarily BDB MySQL)

3 Graph Model

Graph database it stores the value of nodes, edges and properties. There are some general graph database are available that stores any graph and some special kinds of graph database are also available like triple store and network database. In network database it uses edges and nodes to represent and store the data. Graph database is faster when compare to the relational database it map more directly to the structure of object-oriented applications And they successfully implemented in.

- Social networking
- Represent the real world

Is the one of the best NOSQL type to make mind mapping from that we can able mapping the brain and forming a fuzzy based intelligence.

EXAMPLE

```
Node firstNode = graphDb.createNode();
Node secondNode = graphDb.createNode();
Relationship relationship = firstNode.createRelationshipTo(secondNode,
MyRelationshipTypes.KNOWS );
```

eg. Neo4j

```
firstNode.setProperty( "message", "Arun, " );
secondNode.setProperty( "message", "Raju" );
relationship.setProperty( "message", " son" );
```

The graph will look like this:

```
(firstNode )---KNOWS--->(secondNode)
```

Printing information from the graph:

```
System.out.print( firstNode.getProperty( "message txt" ) );
System.out.print( relationship.getProperty( "message txt" ) );
System.out.print( secondNode.getProperty( "messagetxt" ) );
```

eg. Neo4j

Printing will result in:

```
Arun son Raju
```

3.1 How far the Graph Model Helps To Effective Data Retrieval in Robots

In Triple store database it can store triple meaningful data but in the humanoid we need to process with large grouping of data in such cases graph database provides more convenient way of approach in storing the data's in the nodes and its properties describes how it can be search using the queries and the edges shows the relationship among the nodes so that it can form any group based on the query data. As we discussed earlier the humanoid queries are chained so that graph approach provides quicker access without join operation, the humanoid can apply any type of searching mechanism in finding the data's. the results of graph database are different when compare to the other database here we can provide Provenance query answerability this approach is useful when there is a situation in need of making comparative analysis this method can be applied, in this method we can suggest the answer of second question by analyzing the first result. This concept is helpful in decision making in humanoid.

4 Our Proposal System NOSQL On Humanoid Brain

4.1 Humanoid Functionality

Robot electronic system it can't recognize human speech and image. it can repose only to the binary number. Generally the binary digits are eight bits in length. In robot instructions are spited into many pieces and stored in many places because the instruction are in a form of chains, if one instruction starts it continued by another instruction, to explain briefly the robot parts are divided into pieces, for example take

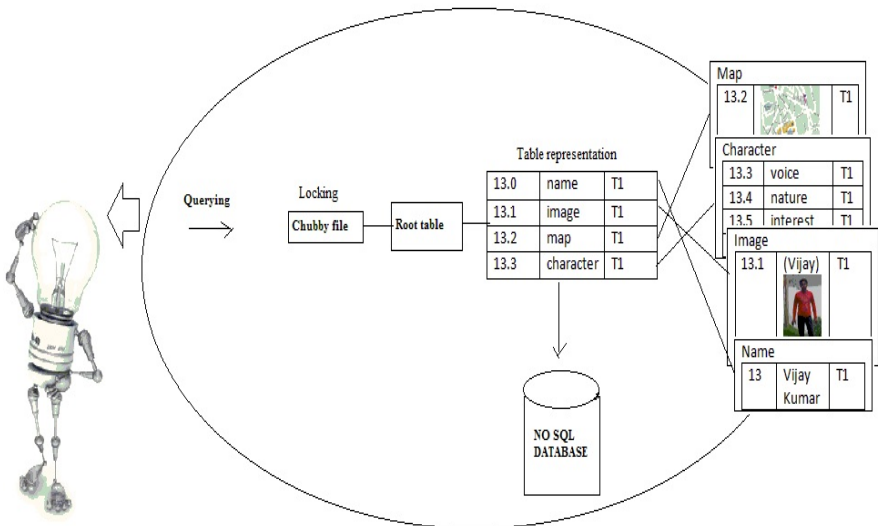


Fig. 2. Implementation of Nosqlon Humanoid Artificial Brain

a hand it has following parts modified spiral joint, revolute joints, spherical joint, phalanges, knuckles etc. if we want to take an object or a particle we need to move all these parts to perform the given task and we need an instructions need to be frame it. If we pass multiple pieces of Instructions to move a hand itself we need a multiprocessor system to process all those instructions, assume that the robot has various parts that made to work to perform a task. The situation is seemed to be more complex, to resolve this conflict we going to provide a solution for this, instead of passing group of instructions, the instructions are passed based on the task it will first call one instruction and it address another instruction then it continues until all instructions are passed if we want to follow this mechanism we need an effective database to handle with this much of instruction.

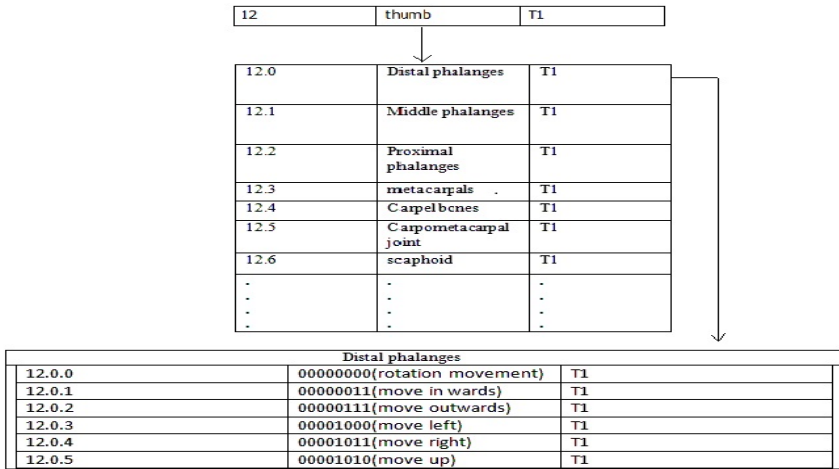


Fig. 3. NOSQL DB Model for Robotic instruction

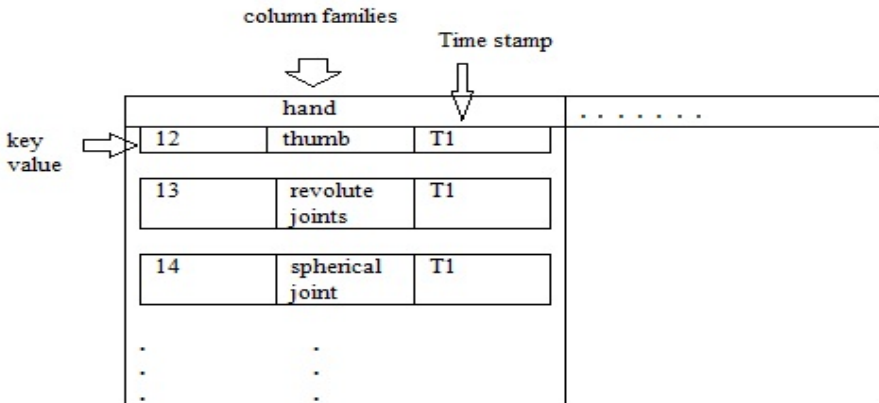


Fig. 4. NOSQL BD Model for Robotic instruction Information: Grouping of instruction for thumb (back node)

4.2 More Memory Using NOSQL

Nowadays NOSQL is the popular non-relational database it can handle with terabytes of data, it has an time stamp mechanism so that it queries current data without need of any special query to retrieve the latest information, it is very helpful in the robots because the robots will scan and updates the data's regularly so that the time stamp mechanism helps to reduce the processing time. In NOSQL database it has a column family's concept the Column keys are grouped into sets called column families, which form the basic unit of access control. All data stored in a column family is usually of the same type. By using this concept we spilt the instructions based on the task and the portion need to be moved, so that we can easily organize the data's (instructions) in column families, so that with the help of one object we can easily refer the entire object based on the task.

5 Map Reduces in Robotics

Map reduce is the framework for processing the large problems, it is need in robots because robots can scan large image and it will try to stored it in database at that time the Database finds difficulty to break the image into pieces and store the respected set. If suppose the robot made a query to Match the current scanning image with the database at that time the database find difficulty to combine the data that are stored, so that it need to focus more on queries to remove all this drawbacks map reduce was introduced it helps to divide the problems into many pieces and it given to the several worker node.

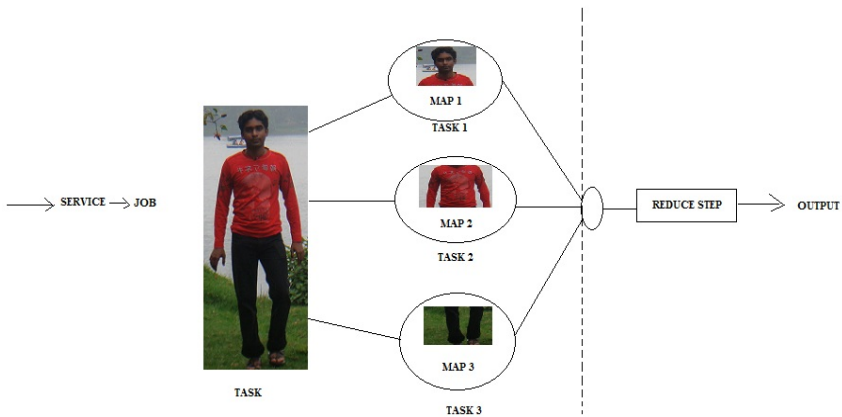


Fig. 5. Map Reduce Proposal System Model On Robotics

5.1 Master Node

The master node takes the input and it assigns the work to the clients.

5.2 Reduce Step

In the reduce step the result are combined and given to the master node from that figure 5 you can easily get idea about that.

5.3 Importance

Above will explain you about the data management and retrieve capacity of NOSQL but this Map Reduce give a massive performance in terms of input and output process from NOSQL.

6 Advantages of Our NOSQL Implementation on Robotics

NOSQL is schema-free database's so it is easy to implement and maintain, it can scale up and down, these database's are replicated to avoid fault-tolerant and can be partitioned if it scales large, the data are easily distributed to the database's, it can process large amount data within a short period of time, it supports specific problem/situation that are no need to think in terms of relations but in terms given in a situation (e.g. documents, nodes,...) in most cases it is freely available because in most of the products are open-source.

First we don't go with huge data server so it gives consume cost and it also have ability to consume power, from this proposal you can able to achieve superfast because we can split the input and Output easily using this map reduce concept from that we easily go with cloud computing.

Then the concept time stamp we easily update our robots without any redundancy of data so we easily avoid garbage collection and give automation to unwanted information deletion and the important thing is the updating process happens immediately when you do this. E.g. like updating the cricket score after updating it immediately publishes that info and replace the old score. It has ability to handle billions of objects so we improve the vision intelligence and hearing intelligence so it is the major step to produce humanoid with very high sensitive intelligence. E.g. the database used for Amazon S3, which as of March 2010 was hosting 102 billion objects.

7 Data Retrievals and Meaningful Schemas Issues in Database

In the relational database design the table schemas are framed by applying the 3 normal forms so that it did not have data redundancy issues while retrieving the data sets and the each table contains meaningful schemas. But in NOSQL it did not focus on what kind of data it stores and how the design should be it entirely focus on how it can be stored effectively. In NOSQL the data sets are divided into multiple pieces of information and each carries its own time stamps based on the arrivals. By using the map reduce it stores the data in a convenient way then the question arises how the data can be retrieved? By using the column families it gives what kind of data it stores based on the some relations and through the time stamp mechanism we obtain the current data. NOSQL supports garbage collections removal so that the old time

stamped messages can be removed easily. We already said that humanoid will use its instructions based on fuzzy sets so that instructions can be linked in a chained manner. The chained instructions is given to the map it splits its job to the respective workers in reduce step so that we can process chained queries in such a way that the humanoid can think.

8 Conclusions

In future our database concept can be major part in making robotics it brings an effective way in making the robots to think. This database concept can be applied in distributed environment so that the data can be maintained in separate place and the updates can be done through networking. From NOSQL we can easily achieve the new era of humanoids.

Acknowledgments. The authors are grateful to Dr. S. Kannan Advisor of project6thsense and to all team members of project6thsense for the making of our entire research projects and for your valuable feedbacks.

References

1. Abadi, D.J., Madden, S.R., Ferreira, M.C.: Integrating compression and execution in columnoriented database systems. In: Proc. of SIGMOD (2006)
2. Ailamaki, A., Dewitt, D.J., Hill, M.D., Skounakis, M.: Weaving relations for cache performance. The VLDB Journal, 169–180 (2001)
3. Banga, G., Druschel, P., Mogul, J.C.: Resource containers: A new facility for resource management in server systems. In: Proc. of the 3rd OSDI, pp. 45–58 (February 1999)
4. Baru, C.K., Fecteau, G., Goyal, A., Hsiao, H., Jhingran, A., Padmanabhan, S., Copeland, G.P., Wilson, W.G.: DB2 parallel edition. IBM Systems Journal 34(2), 292–322 (1995)
5. Bavier, A., Bowman, M., Chun, B., Culler, D., Karlin, S., Peterson, L., Roscoe, T., Spalink, T., Wawrzoniak, M.: Operating system support for planetary-scale network services. In: Proc. of the 1st NSDI, pp. 253–266 (March 2004)
6. Bentley, J.L., Mcilroy, M.D.: Data compression using long common strings. In: Data Compression Conference, pp. 287–295 (1999)
7. Bloom, B.H.: Space/time trade-offs in hash coding with allowable errors. CACM 13(7), 422–426 (1970)
8. Burrows, M.: The Chubby lock service for loosely coupled distributed systems. In: Proc. of the 7th OSDI (November 2006)
9. Chandra, T., Griesemer, R., Redstone, J.: Paxos made live. An engineering perspective. In: Proc. of PODC (2007)
10. Comer, D.: Ubiquitous B-tree. Computing Surveys 11(2), 121–137 (1979)
11. Copeland, G.P., Alexander, W., Boughter, E.E., Keller, T.W.: Data placement in Bubba. In: Proc. of SIGMOD, pp. 99–108 (1988)
12. Dean, J., Ghemawat, S.: MapReduce: Simplified data processing on large clusters. In: Proc. of the 6th OSDI, pp. 137–150 (December 2004)
13. Dewitt, D., Katz, R., Olken, F., Shapiro, L., Stonebraker, M., Wood, D.: Implementation techniques for main memory database systems. In: Proc. of SIGMOD, pp. 1–8 (June 1984)

14. Dewitt, D.J., Gray, J.: Parallel database systems: The future of high performance database systems. *CACM* 35(6), 85–98 (1992)
15. French, C.D.: One size _ts all database architectures do not work for DSS. In: Proc. of SIGMOD, pp. 449–450 (May 1995)
16. Gawlick, D., Kinkade, D.: Varieties of concurrency control in IMS/VS fast path. *Database Engineering Bulletin* 8(2), 3–10 (1985)
17. Ghemawat, S., Gobiuff, H., Leung, S.-T.: The Google _le system. In: Proc. of the 19th ACM SOSP, pp. 29–43 (December 2003)
18. Gray, J.: Notes on database operating systems. In: Flynn, M.J., Jones, A.K., Opderbeck, H., Randell, B., Wiehle, H.R., Gray, J.N., Lagally, K., Popek, G.J., Saltzer, J.H. (eds.) *Operating Systems*. LNCS, vol. 60, pp. 393–481. Springer, Heidelberg (1978)
19. Greer, R.: Daytona and the fourth-generation language Cymbal. In: Proc. of SIGMOD, pp. 525–526 (1999)
20. Hagmann, R.: Reimplementing the Cedar _le system using logging and group commit. In: Proc. of the 11th SOSP, pp. 155–162 (December 1987)
21. Hartman, J.H., Ousterhout, J.K.: The Zebra striped network _le system. In: Proc. of the 14th SOSP, Asheville, NC, pp. 29–43 (1993)
22. KX.COM: Product page, <http://KX.COM.kx.com/products/database.php>
23. Lamport, L.: The part-time parliament. *ACM TOCS* 16(2), 133–169 (1998)
24. Maccormick, J., Murphy, N., Najork, M., Thekkath, C.A., Zhou, L.: Boxwood: Abstractions as the foundation for storage infrastructure. In: Proc. of the 6th OSDI, pp. 105–120 (December 2004)
25. Mccarthy, J.: Recursive functions of symbolic expressions and their computation by machine. *CACM* 3(4), 184–195 (1960)
26. Robbins, R.J.: Database Fundamentals. Johns Hopkins University, rrobbins@gdb.org
27. NOSQL-databases.org
28. <http://books.couchdb.org/relax/intro/eventual-consistency>
29. <http://www.rackspacecloud.com/blog/2009/11/09/NOSQL-ecosystem/.JonathanEllis>
30. <http://blog.evanweaver.com/articles/2009/07/06/up-and-running-with-cassandra/>
31. <http://horicky.blogspot.com/2009/11/NOSQL-patterns.html>
32. <http://www.slideshare.net/Eweaver/cassandra-presentation-at-NOSQL>
33. Bigtable: A Distributed storage system for structured data, google, Inc. In: OSDI, Map Reduce: Simplified Data Processing on Large Cluster, Google, Inc. (2006)
34. Vijay Kumar, S., Saravanakumar, S.G.: Revealing of NOSQL Secrets. *CiiT Journal* 2(10), 310–314 (2010), doi=<http://www.ciiiresearch.org/dmkeoctober2010.html>
35. Vijay Kumar, S., Saravanakumar, S.G.: Emerging Trends in Robotics and Communication Technologies. In: INTERACT 2010 (December 2010), doi=http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=5706225

Impact of Outliers on Anonymized Categorical Data

K. Venkata Ramana, V. Valli Kumari, and K.V.S.V.N. Raju

Department of Computer Science and Systems Engineering, Andhra University
Visakhapatnam, Andhra Pradesh, India, 5300 03

{kvramana.777, vallikumari, kvsvn.raju}@gmail.com

Abstract. Preserving privacy is indispensable when publishing microdata with sensitive information. Anonymization principles like k -anonymity, l -diversity were developed to protect the sensitive information. An adversary with sufficient background knowledge inferring the individual's sensitive information signifies disclosure of the microdata. None of the above mentioned principles addressed the presence of outliers. Outliers can be classified into two types viz., local and global. This paper proposes a practically feasible distance based algorithm to anonymize the local outliers. Our proposed algorithm is capable of handling both numerical and categorical data. The experimental results of our proposed approach focused to categorical data presented.

Keywords: privacy preserving, outlier's detection, distance based anonymization.

1 Introduction

Privacy is the interest that an individual has in sustaining a 'personal space' free from inference by other people and organizations. In recent years, with rapid growth of database, networking and computing technologies a large amount personal data are collected and stored for the purpose of research, statistical analysis, business analysis and other public benefit purposes. Depending upon the nature of the information, data publisher (data holder) may not be interested in revealing the person specific data records. The data of the individuals can be identified directly by their unique values such as name, ssn etc. So, the data publisher before publishing the table removes these identifying attributes to avoid the privacy breach. But sometimes combination of attributes like zip code, age, sex when combined with external source might derive individual sensitive information. The combination of such attributes was termed as a Quasi-Identifier and the attack caused by this is termed as linkage attack [3].

Techniques like perturbation, sampling, data swapping, substituting, randomization, generalization, fuzzyfication, global recoding, local recoding [1][4][6][7][9][10] were applied on the microdata to avoid the disclosure of confidential information. Data generalization is more popular technique being used to avoid the linking attacks mostly when more specific values was replaced with less specific value such that the semantics are preserved. Several principles such as k -anonymity, l -diversity, m -invariance, (α, k) -anonymity were proposed to achieve privacy with the help of aforementioned methods. But the aforementioned principles

do not consider the impact of outliers on the anonymized data. Hui Wang et.al proposed a method called plain k-anonymity which provides identifying the outliers and solving the outliers' problem [2]. However it focuses on the numerical attribute only.

1.1 Motivation

The data elements are said to be outliers if they are considerably dissimilar from the remaining data elements. For instance, consider the hospital dataset as shown in the table 1(a) where name is the identifying attribute, age, sex and education are Quasi-Identifiers and disease attribute is considered as sensitive. After applying k-anonymity and l-diversity on table 1(a) the resultant generalized table is shown in table 2(a). Even after generalization there is a possibility that the data contains outliers. The adversary with some background knowledge can derive individual sensitive values due to the presence of outliers.

Example 1: Suppose the adversary has some background knowledge about Bob's quasi-identifier group and sensitive information. Bob's quasi-identifier group is age [20-30], gender [M, F] and he is suffering with long term viral disease. The sensitive information of Bob's quasi-identifier group contains gastric ulcer, gastritis and HIV. From this knowledge the adversary concludes that bob is infected with HIV. This shows outliers described with respect to QID group called local outliers [2].

Example 2: Let us assume that the adversary knows the distribution of the sensitive attribute domain as shown in the fig 1. The hierarchies of sensitive attribute disease and its domain values are {HIV, Gastric Ulcer, Gastritis, Stomach Cancer, Malaria, Tuberculosis, Hepatitis, Colitis}. Since the adversary knows the entire domain knowledge and endometriosis is a different type of disease among the domain. With this we can conclude that the person is female (F). When we link with external source we can identify that the individual is Jessica. We term these kind of sensitive values as global outliers [2].

Table 1. (a): Raw Micro data: Bob and Jessica are two outliers

TupleID	Name	Age	Sex	Education	Disease
1	Irma	23	F	Arts	<i>Gastric ulcer</i>
2.	Ahmed	25	M	Sciences	<i>Gastritis</i>
3.	Bob	27	M	Masters	<i>HIV</i>
4.	Ann	31	F	10 th	<i>Gastritis</i>
5.	Dick	33	F	9 th	<i>Stomach cancer</i>
6.	Diana	38	F	10 th	<i>Malaria</i>
7.	John	35	M	11 th	<i>Tuberculosis</i>
8.	Frank	48	F	Sciences	<i>HIV</i>
9.	Cary	47	M	Doctorate	<i>Hepatitis</i>
10	Smith	46	M	Sciences	<i>Stomach cancer</i>
11	Jessica	50	F	12 th	<i>Endometriosis</i>

Table 1. (b): 3-Anonymity 3-diverse

Tuple ID	QID_i Group	Age	Sex	Education	Disease
1		[20-30]	Person	University	Gastric ulcer
2.	QID ₁	[20-30]	Person	University	Gastritis
3.		[20-30]	Person	University	HIV
4		[31-40]	Person	Juniorsec	Gastritis
5.	QID ₂	[31-40]	Person	Juniorsec	Stomach cancer
6		[31-40]	Person	Juniorsec	Malaria
7.		[31-40]	Person	Juniorsec	Tuberculosis
8.	QID ₃	[41-50]	Person	AnyEducation	HIV
9.		[41-50]	Person	AnyEducation	Hepatitis
10.		[41-50]	Person	AnyEducation	Stomach cancer
11.		[41-50]	Person	AnyEducation	Endometriosis

Table 2. 3-Anonymity 3-diverse

Tuple ID	QID Group	Age	Sex	Education	Disease
1		[20-50]	Person	University	Gastric ulcer
2.	QID ₁	[20-50]	Person	University	Gastritis
10.		[20-50]	Person	University	Stomach cancer
4		[31-40]	Person	Juniorsec	Gastritis
5.	QID ₂	[31-40]	Person	Juniorsec	Stomach cancer
6		[31-40]	Person	Juniorsec	Malaria
7.		[31-40]	Person	Juniorsec	Tuberculosis
8.	QID ₃	[20-50]	Person	AnyEducation	HIV
9.		[20-50]	Person	AnyEducation	Hepatitis
3.		[20-50]	Person	AnyEducation	HIV

2 Related Work

Sweeney and Samarati proposed k-anonymity model [1] [3], partitioning the domain of each quasi-identifier attribute into intervals and replacing the values in the attribute with the intervals to which the values belong to using a concept tree. Then the records in the table are naturally grouped by the same intervals of the quasi-identifiers, if the sizes of all groups are at least k . However k -anonymity does not focus on sensitive information. To solve this l -diversity model [5], was proposed.

l -diversity states that for each QID group there must be at least l “well represented” sensitive values. Other flavors of k-anonymity like (α, k) -anonymity and (c, k) safety were proposed. But none of these techniques considers outliers.

R.C.Wong et.al proposed local and global recoding principles [7]. They describe the relation between quasi-identifiers and single sensitive attribute. Hui et.al [2] discussed distinguishability based attack due to local and global outliers. They proposed method called plain k-anonymity in which every individual tuple in a QI

group must contain at least k distinct sensitive values. Their method restricts to numerical sensitive values.

This paper examines categorical outliers and their impact on privacy in anonymized data.

3 Preliminaries

In order to be familiar with the concepts that were dealt in this paper we discuss a few insights on the outliers and basic definitions in this section.

3.1 Outliers in Anonymized Data

Outlier is the one that appears to be deviated markedly from other members of the sample in which it is present. As stated in the earlier sections outliers are classified as global and local outliers.

Let an anonymized QI - group be QIG and the sensitive data object $o \in QIG$. If at least p percentage of sensitive values lie at the distance more than d from o , we term as local outliers and if the distance between o and other domain values is greater than 1 we term as global outliers.

$$\begin{aligned} \text{Distance}(o, s) \geq d & \quad \text{Local Outlier} \\ > 1 & \quad \text{Global Outlier} \quad \text{where } s \in QIG[S] \text{ and } d \in [0,1] \end{aligned}$$

Outliers can be detected using methods like distance based, density based, depth based, deviation based and statistical based outliers [11][12][13][14][15]. In this paper, we adopted distance based outlier measure for detecting outliers.

Definition1. (Distance based outlier):

An outlier is a data object ‘ o ’ in a dataset D with some percentage of the objects in D having a distance of more than d away from itself. We consider Euclidian distance for continuous attributes and hierarchical distance for categorical attributes as explained in section 3.3.

3.2 Anonymization Frame Work

Let T be the microdata table that needs to be published. The table contains collection of tuples from domain $D = A \times B \times C = A_1 \times A_2 \times \dots \times A_i \times B_1 \times B_2 \times \dots \times B_q \times C_1 \times C_2 \times \dots \times C_s$ were, A_1, A_2, \dots, A_i are identifying attributes, B_1, B_2, \dots, B_q are quasi identifiers and C_1, C_2, \dots, C_s are sensitive attributes. For each tuple $t \in T$, the attribute value is denoted as $t[A]$.

Definition 2. (Generalization):

The domain D of an attribute A_i is split into a set of disjoint partitions (subsets), such that for any tuple $t \in D_A$ and any attribute A of t , $I_{t[A_i]} \preceq I_{f(t[A_i])}$, where $I_{t[A_i]}, I_{f(t[A_i])}$ are the intervals of the $t[A]$ and $f(t[A_i])$, \preceq is the binary relation and f is the generalization function.

Definition 3 (K-Anonymity):

The table is said to be k-anonymous if every combination of quasi identifier attribute values in a generalized table occurs k or more times [3].

Definition 4 (QI- group):

The microdata domain D splitting into set of equivalence classes, such that each tuple of quasi identifier in D belongs to exactly one class. All the tuples of each class are identical quasi identifier values. We refer to these classes QI- groups denoted as $QID_1, QID_2... QID_m$,

$$\bigcup_{i=1}^m QID_i = D \text{ and } QID_p \cap QID_q = \emptyset, \forall p, q \ 1 \leq p \neq q \leq m.$$

The equivalence classes are specified as domain intervals for numerical attributes and higher level value of taxonomy tree for the categorical attribute.

Definition 5 (Anonymized QI - Group):

An anonymized dataset D^* contains anonymized QID – groups and sensitive values. The anonymized QID – group (QIG) = $QID_i \cup S_i$, Here QID_i is the i^{th} Quasi identifier group (QID) and S_i is the set of sensitive values which contain in i^{th} QID .

For example in table 1 Bob’s QIG is [20-30] [person], [University] \cup {HIV, Gastric ulcer, Gastritis}

Definition 6 (l- diversity):

An anonymized quasi identifier group (QIG) is said to be l-diverse if it contains at least ‘l’ distinct sensitive values [5].

Gastric ulcer	Gastritis	HIV	Stomach Cancer	Malaria	Tuberculosis	Hepatitis
1	2	3	4	5	6	7

Fig. 1. Indexing of categorical attributes

3.3 Metrics of Outliers

3.3.1 Hierarchical Distance: Let v_1 and v_2 be the two categorical attribute values. The distance between v_1 and v_2 is minimum level to which these two values are generalized to the same value in the taxonomy tree.

Distance $(v_1, v_2) = \frac{level(v_1, v_2)}{H}$ here level (v_1, v_2) is the lowest common ancestor of values in the taxonomy tree and H is height of tree.

Hierarchical distance always lies between [0, 1]. If the distance is 0 it signifies that both values are same. If the values have the root node of the tree as the common ancestor its distance is 1. For example consider disease domain hierarchical tree as shown in Fig 2. The distance between “HIV” and “Malaria” is 2/3, distance between “Stomach cancer” and “HIV” is 1 and distance between “gastric ulcer” and “gastric ulcer” is 0.

3.3.2 Set Distance: If X and Y are the two non-empty sets the distance between two sets X and Y is the average distance of each point of the first set to each point of the second set.

$$Sdis(X,Y) = \frac{\sum_{x_i \in X, y_j \in Y} distance(x_i, y_j)}{|X| * |Y|}, \quad \forall x_i \in X, \quad y_j \in Y$$

Here |X| and |Y| are the number of elements in the first set and second set. For example: the set $X_1 = \{\text{gastric ulcer, gastritis, HIV}\}$ and $X_2 = \{\text{HIV, Hepatitis}\}$. The set distance between X_1 and X_2 is $(1+1+1+0+1/3)/6 = 0.7221$.

1	0						
2	1/3	0					
3	1	1	0				
4	1/3	1/3	1	0			
5	1	1	2/3	1	0		
6	1	1	2/3	1	1/3	0	
7	1	1	1/3	1	2/3	2/3	0
	1	2	3	4	5	6	7

Fig. 2. Distance matrix for the categorical attributes

4 Algorithm

Given microdata D, our goal is to split data into set of equivalence classes which satisfies *k*- anonymity, *l*- diversity and does not have outliers, so that the adversary cannot infer the any knowledge from anonymized *QI*- group. To achieve this goal we propose the frame work. This frame work consists of two phases. In the first phase we find the outliers of each anonymized *QI*-group which is discussed in algorithm 1. In second phase global outliers are eliminated and local outliers are merged to an appropriate anonymized group which is discussed in algorithm 2.

Phase 1: In this phase, initially we find the center of data object value of the each anonymized quasi identifier group (*QIG*). The selection of the center of each *QIG* is based on the frequency of data element within that group. If all values are distinct we randomly pick an element that could be a mode. Using the center object, all the hierarchical distances of remaining elements in the *QIG* are determined. If the distances are more than defined threshold we treat that to be a local outlier. Once the local outliers are figured out we maintain a local-outlier list. If the hierarchical distance is more than 1, the data elements are called global outliers and they are maintained in the global outlier list.

For instance consider the first *QIG* in table 1(a) that contain the elements like {gastric ulcer, gastric, HIV} and defined threshold 0.5. If within the set, all elements are distinct then select first element as center object of the group. Therefore, “HIV” is local outlier because the distance between “HIV” and “Gastric ulcer” is 0.67 (more than threshold). In the table 1(b), third *QIG* contain four elements {HIV, Hepatitis, Stomach cancer, Endometriosis}. This set contains both local outlier (stomach cancer) and global outlier (Endometriosis) with respect to “HIV” or “Hepatitis”. The distance matrix is shown in the figure 2.

Phase 2: This phase consists of four steps as shown in algorithm 2. The first step is to eliminate the global outliers from the outliers list, then outliers list has only local outliers. The second step is to form sub lists from list of local outliers. The element of each sub list is close to remaining elements i.e distance between any two elements in the sub list is less than threshold. For example the list contains elements like {HIV, Hepatitis, Gastric ulcer, Stomach cancer}. This set (list) is split into two sub lists {HIV, Hepatitis},{Gastric ulcer, Stomach cancer}.The third step performs merging each sub list to appropriate QIG^l which were obtained in phase 1. The basic metric to merge the sub list to QIG^l is set distance (see 3.3.2 of section 3). If the set distance between the QIG^l and sub list is less than the threshold we combine both. For example $QIG_1^l = \{\text{Stomach cancer, Gastric ulcer}\}$, $QIG_2^l = \{\text{gastric, stomach cancer, malaria, tuberculosis}\}$, outlier sub list $l_1 = \{\text{Gastric, Gastric ulcer}\}$ and threshold $\gamma = 0.45$. The distance between (QIG_1^l, l_1) is 0.333 and (QIG_2^l, l_1) is 0.625. Hence, we merge l_1 to QIG_1^l . Finally in fourth step we rearrange all the quasi identifier values of the QIG^l .

Algorithm 1. Finding the Outliers

Input: Set of Anonymized Quasi identifier groups $G = \{g_1, g_2, g_3, \dots, g_p\}$
 Domain Hierarchy distance matrix of sensitive attribute.

Output: Set of outliers list L

Method: Detection of outliers of anonymized quasi identifier group as follows

1. **Begin**
 2. Finding the center object value of sensitive attribute in each anonymized Quasi Identifier group $O = \{o_1, o_2, o_3, \dots, o_p\}$
 3. Loutliers = { }, Goutliers = { }, initialize the threshold δ .
 4. **for** each g_i in G **do**
 5. **for** each sensitive value t_{js} of the tuple t_j in g_i
 6. **if** (distance(o_i, t_{js}) > 1)
 7. Goutliers = Goutliers \cup { t_{js} }
 8. **if** (distance(o_i, t_{js}) > δ)
 9. Loutliers = Loutliers \cup { t_{js} }
 10. **end for**
 11. **end for**
 12. L = Loutliers \cup Goutliers
 13. $G^l = G - L$
 14. **end Begin**
-

Algorithm 2. Regrouping the Outliers

Input: Set of Anonymized Quasi identifier groups $G^1 = \{g_1^1, g_2^1, g_3^1, \dots, g_p^1\}$ obtained from Algorithm 1.
List of Local outliers, List of Global outliers.

Output: Anonymized quasi identifier group without outliers.

Method:

1. **Begin**
- Step 1: Remove all the global outliers**
2. for all tuple t that is sensitive-attribute outlier $t_{is} \in G_{outliers}$
- do
- Remove t
- Step 2: Arrangement of all of local outliers set into sub sets.**
3. Local outlier set is splitting into sub sets
 i.e. $L_{outlier} = \{l_1, l_2, l_3, \dots\}$
4. **for** each subset l_i in $L_{outlier}$
5. distance(t_{js}, t_{ks}) $< \delta$
6. **end for**
7. **Step 3: Merging each sub set of $L_{outlier}$ to appropriate g_i of G^1 based on setdistance .**
8. **for** each l_i in $L_{outlier}$ do
9. **for** each g_j^1 in G^1 do
10. **if** ($S_{dis}(l_i, g_j^1) < \gamma$)
11. Merge(g_j^1, l_i)
12. **end if**
13. break;
14. **end for**
15. **end for**
16. **Step 4: Re- arranging the quasi identifiers attributes.**
17. **end begin**

5 Performance Analysis

In this section, we determine the complexity of our proposed approach. Let n be the number of tuples and p be the anonymized quasi identifier groups. The complexity of first algorithm is readily split into two parts. The first part of the algorithm 1 is to find the centers of each QIG for determining the local and global outliers. The complexity of finding the centers of each QIG is $O(n)$ and the average case complexity of detecting the local and global outliers is $O(np)$. Hence overall complexity of algorithm1 is $O(np+n)$. In Algorithm 2, we split the local outliers into subsets and merge these local outlier sets to QIG's. Let m be the number of local outliers divided into s number of sets. The complexity of splitting the local outliers into subsets will be $O(m/s)$ and complexity of merging local outlier sets to QIG's is $O(mp/s)$. Hence, the overall complexity of second algorithm is $O(mp/s) + O(m/s)$.

6 Experiment Analysis

Experiments are performed on Intel Core2Duo @ 2.93 GHz on Netbeans platform. The experimentation was conducted on the Adult dataset available at UCI Machine Learning Repository [8]. The dataset consists of age, sex and education attributes as Quasi Identifiers. Disease attribute was taken as sensitive attribute. We assumed that each equivalence class is of size 1,000 tuples. This can be changed if necessary. Fig 3 shows how the outliers vary with increasing of data size at distance threshold 0.6. Fig 4 shows the percentage of tuples with increase of distance threshold δ considering dataset of size 10,000.

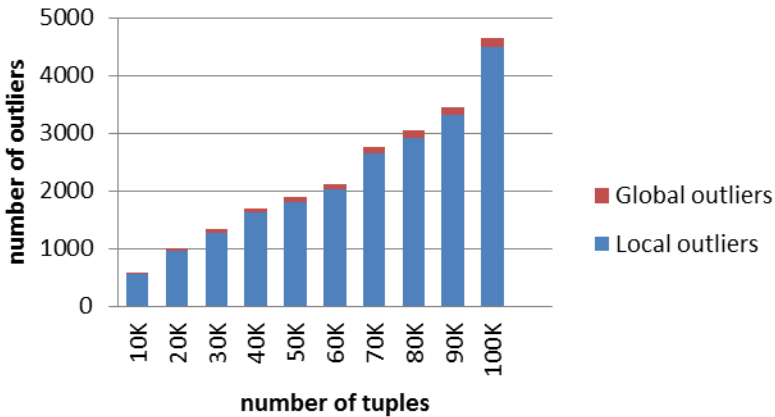


Fig. 3. Detection of local and Global outliers at $\delta = 0.6$

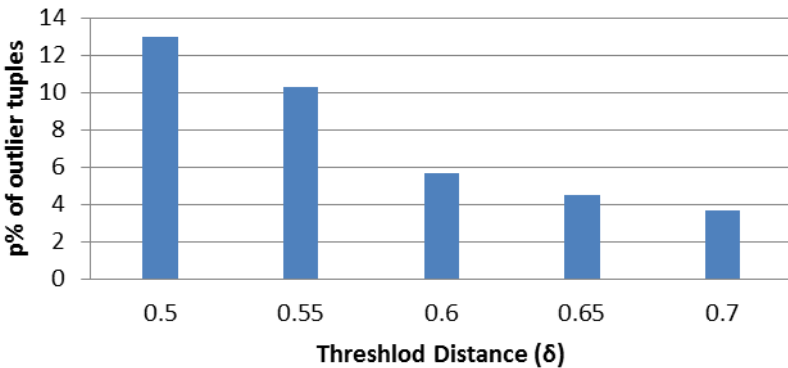


Fig. 4. Percentage of outliers varying with δ

6 Conclusions

Privacy preserving plays a major role when publishing person specific data for mining or survey purposes. The adversary can infer particular individual if he possesses some background knowledge. This paper proposes a distanced based approach for detecting outliers during the anonymization process. With our proposed algorithms we rearrange the local outliers into corresponding QI groups based on distance metric and remove the global outliers. Our algorithm is capable of anonymizing both numerical and categorical data.

References

1. Sweeney, L.: Achieving k-anonymity privacy protection using generalization and suppression. *International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems* 10(5), 571–588 (2002)
2. Wang, H., Liu, R.: Hiding Distinguished Ones into Crowd: Privacy-Preserving Publishing Data with Outliers. In: *The 12th International Conference on Extending Database Technology (EDBT)*, Saint-Petersburg, Russia, March 23–26 (2009)
3. Sweeney, L.: k-anonymity: a model for protecting privacy. *International Journal on Uncertainty, Fuzziness and Knowledge-Based Systems* 10(5), 557–570 (2002)
4. Estivill-Castro, V., Brankovic, L.: Data swapping: Balancing privacy against precision in mining for logic rules. In: Mohania, M., Tjoa, A.M. (eds.) *DaWaK 1999*. LNCS, vol. 1676, pp. 389–398. Springer, Heidelberg (1999)
5. Machanavajjhala, A., Gehrke, J., Kifer, D.: l-diversity: Privacy beyond k-anonymity. In: *ICDE*, p. 24 (2006)
6. Agrawal, R., Srikant, R.: Privacy preserving data mining. In: *Proc. of the 2000 ACM SIGMOD International Conference on Management of Data*, Dallas, Texas, pp. 439–450 (May 2000)
7. Wong, R.C.W., Fu, A.W.C., Wang, K., Pei, J.: Minimality attack in privacy preserving data publishing. In: *VLDB*, pp. 543–554 (2007)
8. Newman, D.J., Hettich, S., Blake, C.L., Merz, C.J.: *UCI Repository of Machine Learning Databases*. University of California, Irvine (1998), <http://www.ics.uci.edu/mllearn/MLRepository.html>
9. Xu, J., Wang, W., Pei, J., Wang, X., Shi, B., Fu, A.: Utility-based anonymization using local recoding. In: *SIGKDD* (2006)
10. Valli Kumari, V., Srinivasa Rao, S., Raju, K.V.S.V.N., Ramana, K.V., Avadhani, B.V.S.: Fuzzy based approach for privacy preserving publication of data. *IJCSNS International Journal of Computer Science and Network Security* 8(1), 115–121 (2008)
11. Knorr, E.M., Ng, R.T., Tucakov, V.: Distance-based outliers: algorithms and applications. *VLDB Journal* 8(3–4), 237–253 (2000)
12. Ramaswamy, S., Rastogi, R., Shim, K.: Efficient algorithms for mining outliers from large data sets. In: *SIGMOD* (2000)
13. Breunig, M.M., Kriegel, H., Ng, R.T., Sander, J.: LOF: Identifying Density-Based Local Outliers. In: *SIGMOD* (2000)
14. Barnett, V., Lewis, T.: *Outliers in Statistic Data*. John Wiley’s Publisher, NY (1994)
15. Arning, A., Agrawal, R., Raghavan, P.: A Linear Method for Deviation Detection in Large Databases. In: *2nd International Conference on Knowledge Discovery and Data Mining Proceedings*, pp. 164–169 (1996)

Privacy in Mobile Ad Hoc Networks

B. Jhansi Vazram¹, V. Valli Kumari², and J.V.R. Murthy³

¹ Dept. of CSE, Narasaraopeta Engg. College, Andhra Pradesh, India-522601

² Dept. of CS &SE., Andhra University, Andhra Pradesh, India – 530003

³ Dept. of CSE, J N T University, Kakinada, Andhra Pradesh, India – 533003
{jhansi.bolla,vallikumari,mjonnalagedda}@gmail.com

Abstract. In mobile adhoc networks, generating and maintaining anonymity for any adhoc node is challenging because of the node mobility, dynamic network topology, cooperative nature of the network and broadcast nature of the communication media. Anonymity is provided to protect the communication, by hiding the participants as well as the message contents. Existing techniques based on cryptosystem and broadcasting cannot be easily adapted to MANET because of their extensive cryptographic computation and/or large communication overhead. In this paper, we first propose an unconditionally secure privacy preserving message authentication scheme (PPMAS) which uses Modified New variant ElGamal signature Scheme (MNES). Secondly we proposed privacy preserving communication protocol for MANET based on dynamic generation of pseudonyms, which are used in place of real nodes to provide anonymity.

Keywords: Network security, anonymity, mobile adhoc networks.

1 Introduction

A mobile ad hoc network (MANET) comprises of a set of wireless devices that can move around freely and cooperate in relaying packets on behalf of one another. A MANET does not require a fixed infrastructure or centralized administration. Distant mobile nodes communicate through multi hop paths, as they have limited transmission range. Their ease of deployment makes MANETs an attractive choice for a variety of applications like battleground communications, disaster recovery efforts, communication among a group of islands or ships, conferencing without the support of a wired infrastructure, and interactive information sharing. In MANETs, mobile nodes cooperate to forward data on behalf of each other. Typical protocols used for self-organizing and routing in these networks expose the node identifiers (network and link layer addresses), neighbors, and the end-points of communication. Some modes of operation further mandate that the nodes freely divulge their physical location. In short, nodes must advertise a profile of their online presence to participate in the MANETs, which is highly undesirable.

Both military and civilian MANETs may find the mandated exposure of information unacceptable, a node should be able to keep its identity, its location and its correspondents private, i.e., remain anonymous [6], [4]. Any solution providing

anonymity must overcome the broadcast nature of wireless environments (which enables eavesdropping) and operate under often tight resource constraints, unlike wired networks. Simple solutions like packet encryption are also largely ineffective because of ease of traffic analysis over a broadcast media. Hence, supporting privacy in MANETs is enormously challenging.

Outline of the paper: Section 2 presents the related work done. Section 3 gives overview of the proposed privacy preserving unconditionally secure message authentication scheme. Section 4 proposes a privacy preserving communication protocol. Section 5 discusses security analysis. Section 6 gives the performance analysis and finally we conclude in section 7.

2 Related Work

Unlinkability of an entity with a message or an action performed by it means that an adversary with enough information is unable to identify the identity of the entity, given the message or an action performed by it. Unlinkability makes anonymity possible. Privacy of both the source and the destination has to be protected in MANETs. It is also desirable that the attacker should not be able to derive the fact that source and destination nodes are communicating. We define a set of objects called anonymous set (AS) to see that a particular object is unidentifiable. Our proposed work, like any other signature schemes consists of two algorithms, generation and verification. With generation algorithm, given the message m and the public keys of anonymous set (AS), a sender from AS, with her own private key, can generate an anonymous message $s(m)$. The verification algorithm, given the message m and anonymous message $s(m)$, is used to verify whether $s(m)$ is generated by a member in the AS. The security requirements for our method are sender anonymity and Unforgeability.

2.1 Modified New variant ElGamal signature scheme (MNES)

Definition 1: Based on the New variant ElGamal signature scheme [2], we propose a modified new variant ElGamal signature scheme(MNES), which consists of the following 3 algorithms:

- i) *Key generation algorithm* : Let p be a large prime, α be a generator of Z^* . Both p and α are made public. For a random private key $x \in Z_p^*$. The public key y is computed from $y = \alpha^x \text{ mod } p$.
- ii) *Signature algorithm:* The MNES can also have many variants. For the purpose of efficiency, we will describe the variant, called optimal scheme. To sign a message m , one chooses a random $k, l \in Z_{p-1}^*$, then computes the exponentiation $r = \alpha^k \text{ mod } p$, $s = \alpha^l \text{ mod } p$ and $q = \alpha^h \text{ mod } p$ and solves w from

$$w = k + l + h + xr + ks + lq \text{ mod } (p - 1) \tag{1}$$

where h is a one way hash function .The signature of the message m is defined as the triple (r, s, w) .

- iii) *Verification algorithm* :

The verifier checks the signature equation $\alpha^w = rsqy^r r^s s^q \text{ mod } p, h=h(m,rs)$. If the equality holds true, then the verifier accepts the signature and rejects otherwise.

The existing anonymous communication protocols are largely stemmed from either mixnet [1] or DC-net [3]. Moler presented a secure public-key encryption algorithm for mixnet [8]. This algorithm has been adopted by Mixminion [7]. However, since mixnet like protocols rely on the statistical properties of background traffic, they cannot provide provable anonymity.

Recently [5][10] message sender anonymity based on ring signatures was introduced. This method provides an assurance to the sender that the generated message has source anonymous signature along with content authenticity, while hiding the message sender's real identity.

In this paper, we first propose an unconditionally secure privacy preserving message authentication scheme (PPMAS) based on the modified new variant ElGamal signature scheme. This is because the original ElGamal signature scheme is existentially forgeable with a generic message attack [12, 11]. While the modified ElGamal signature (MES) scheme [5] is secure against no-message attack and adaptive chosen message attack in the random oracle model [9], it cannot be used for more than one message. The modified new variant ElGamal signature scheme (MNES) is almost very similar to MES, and also [2] we can transmit more than one message without changing the secret exponents.

3 Unconditionally Secure Privacy Preserving MAC (PPMAS):

In this section, we propose an efficient privacy preserving unconditionally secure message authentication scheme (PPMAS). The main idea is that for each message m to be released, the sending node generates a privacy preserving message authentication for the message m . The generation is based on the MNES scheme. Unlike ring signatures, which require computing a forgery signature for each member in the AS separately, our scheme only requires three steps to generate the entire PPMAS. This scheme links all non-senders and the message sender to the PPMAS alike. In addition, our design enables the PPMAS to be verified through a single equation without individually verifying the signatures.

3.1. The Proposed PPMAS Scheme. Suppose that the message sender (say Alice) wishes to transmit a message m anonymously from her network node to any other node. The AS includes n members, A_1, A_2, \dots, A_n , for example, $\mathcal{s} = \{A_1, A_2, \dots, A_n\}$, where the actual message sender Alice is A_t , for some value $t, 1 \leq t \leq n$.

Let p be a large prime number and α be a primitive element of Z_p^* . Then α is also a generator of Z_p^* . That is $Z_p^* = \langle \alpha \rangle$. Both p and α are made public and shared by all members in \mathcal{s} . Each $A_i \in \mathcal{s}$ has a publickey $y_i = \alpha^{x_i} \text{ mod } p$, where x_i is a randomly selected private key from Z_{p-1}^* . In this paper, we will not distinguish between the node A_i and its public key y_i . Therefore, we also have $\mathcal{s} = \{y_1, y_2, \dots, y_n\}$.

Suppose m is a message to be transmitted. The private key of the message sender Alice is x_t , $1 \leq t \leq n$. To generate an efficient PPMAS for message m , Alice performs the following three steps:

(1) Select a random and pair wise different k_i, l_i for each $1 \leq i \leq n, i \neq t$ and computer $r_i = \alpha^{k_i} \bmod p$, $s_i = \alpha^{l_i} \bmod p$, $q_i = \alpha^{h_i} \bmod p$ where $1 \leq k_i, l_i, q_i < p$.

(2) Choose two integers k, l randomly where $1 \leq k, l < p$ and compute $r_t = \alpha^k \prod_{i \neq t} y_i^{-r_i} \bmod p$, $s_t = \alpha^l \prod_{i \neq t} r_i^{-s_i} \bmod p$ and $q_t = \alpha^h \prod_{i \neq t} s_i^{-h_i} \bmod p$, such that $r_t \neq 1, s_t \neq 1, q_t \neq 1, r_i \neq r_t, s_i \neq s_t, q_i \neq q_t$ for each $i \neq t$.

(3) Compute

$$w = k + s + h + \sum_{i \neq t} k_i + \sum_{i \neq t} l_i + \sum_{i \neq t} h_i + x_t r_t + k_t s_t + l_t q_t \bmod (p - 1).$$

The PPMAS of the message m is defined as

$$s(m) = (m, s, r_1 \dots r_n, s_1 \dots s_n, q_1 \dots q_n, h_1 \dots h_n, w) \quad (2)$$

$$\text{Where } \alpha^w = r_1 \dots r_n s_1 \dots s_n q_1 \dots q_n y_i^{r_1} \dots y_n^{r_n} \dots r_1^{s_1} r_n^{s_n} \dots s_1^{q_1} \dots s_n^{q_n} \bmod p \quad (3)$$

3.2 Verification of PPMAS

A verifier can verify an alleged PPMAS($m, s, r_1 \dots r_n, s_1 \dots s_n, w$) for message m by verifying whether the following equation

$$\alpha^w = r_1 \dots r_n s_1 \dots s_n q_1 \dots q_n y_i^{r_1} \dots y_n^{r_n} r_1^{s_1} \dots r_n^{s_n} \dots s_1^{q_1} \dots s_n^{q_n} \bmod p$$

holds. If (3) holds true, the verifier accepts the PPMAS as valid for message m . Otherwise the verifier rejects the PPMAS.

In fact, if the PPMAS has been correctly generated, then we have

$$\begin{aligned} & \prod_{i=1}^n r_i \prod_{i=1}^n s_i \prod_{i=1}^n q_i \prod_{i=1}^n y_i^{r_i} \prod_{i=1}^n r_i^{s_i} \prod_{i=1}^n s_i^{q_i} \bmod p \\ &= \left(\prod_{i \neq t} r_i \right) r_t \left(\prod_{i \neq t} s_i \right) s_t \left(\prod_{i \neq t} q_i \right) q_t \left(\prod_{i \neq t} y_i^{r_i} \right) y_t^{r_t} \\ & \quad \left(\prod_{i \neq t} r_i^{s_i} \right) r_t^{s_t} \left(\prod_{i \neq t} s_i^{q_i} \right) s_t^{q_t} \bmod p \\ &= \alpha^{\sum_{i \neq t} k_i} \alpha^{\sum_{i \neq t} l_i} \alpha^{\sum_{i \neq t} h_i} \left(\alpha^k \prod_{i \neq t} y_i^{-r_i} \right) \left(\prod_{i \neq t} y_i^{r_i} \right) y_t^{r_t} \\ & \quad \left(\alpha^l \prod_{i \neq t} r_i^{-s_i} \right) \left(\prod_{i \neq t} r_i^{s_i} \right) r_t^{s_t} \left(\alpha^h \prod_{i \neq t} s_i^{-q_i} \right) \left(\prod_{i \neq t} s_i^{q_i} \right) s_t^{q_t} \bmod p \\ &= \alpha^{\sum_{i \neq t} k_i} \alpha^{\sum_{i \neq t} l_i} \alpha^{\sum_{i \neq t} h_i} \alpha^k y_t^{r_t} \alpha^l r_t^{s_t} \alpha^h s_t^{q_t} \bmod p \\ &= \alpha^{\sum_{i \neq t} k_i + \sum_{i \neq t} l_i + \sum_{i \neq t} h_i + k + l + h} \alpha^{x_t r_t} \alpha^{k_t s_t} \alpha^{l_t q_t} \bmod p \\ &= \alpha^{k+l+h + \sum_{i \neq t} k_i + \sum_{i \neq t} l_i + \sum_{i \neq t} h_i + x_t r_t + k_t s_t + l_t q_t} \bmod p \\ &= \alpha^w \bmod p \end{aligned}$$

Therefore, the verifier should always accept the PPMAS if it is correctly generated without being modified.

As a trade-off between computation and transmission, the PPMAS can also be defined as $\mathcal{s}(m) = (m, \mathcal{s}, r_1, \dots, r_n, s_1, \dots, s_n, q_1, \dots, q_n, h_1, \dots, h_n, w)$. In case \mathcal{s} is also clear, it can be eliminated from the PPMAS.

3.3 Security Analysis

In this subsection, we prove that the proposed PPMAS scheme is unconditionally anonymous and provably unforgeable against adaptive chosen-message attack.

3.3.1 Anonymity

In order to prove that the proposed PPMAS is unconditionally anonymous, we have to prove that (i) for anybody other than the members of S , the probability to successfully identify the real sender is $1/n$, and (ii) anybody from \mathcal{s} can generate PPMAS.

3.3.2 Unforgeability

The design of the proposed PPMAS relies on the ElGamal signature scheme. Different levels of security can be achieved by signature schemes. The maximum level of security is a counter to existential forgery under adaptive chosen message attack.

4 The Privacy Preserving Communication Protocol

4.1 Network Assumption

As any physical transmission in a world can be monitored and traced to its origin, it is probably impossible to keep confidential who is communicating to whom by which messages. Our paper addresses the above problem. Assume that our network model similar to that discussed in [5], consists of networks with multiple MANETs, i.e., the participating nodes are divided into set of small groups. The network nodes are categorized into Ordinary node and Special nodes

An ordinary node is one that is unable to communicate directly with the nodes in other MANETS. A special node can be an ordinary node that can also provide message forward services to other MANET nodes. In some peculiar situations e.g.: energy optimization, an ordinary node can be automatically converted to a special node.

Prior to the network deployment, there should be an administrator. The administrator does not take part in routing rather it has the following tasks during the bootstrap of the network.

- i) Determines two groups G_1, G_2 of the same prime order q . We view G_1 as an additive group and G_2 as a multiplicative group.
- ii) Determines bilinear map: $G_1 \times G_1 \rightarrow G_2$, collision resistant cryptographic hash functions H_1 and H_2 where $H_1 : \{0,1\}^* \rightarrow G_1$ mapping from arbitrary length strings to points in G_1 and $H_2 :$

$\{0,1\}^* \rightarrow \{0,1\}^\mu$ mapping from arbitrary length strings to μ bit fixed length output.

- iii) Generates system's secret $\varphi \in Z_q^*$, where $Z_q^* = \{y \mid l \leq y \leq q - l\}$. Any one in the network does not know φ except system administrator.

Thus the parameters $\langle G1, G2, g, H1, H2 \rangle$ are known to the special nodes. System administrator also provides the following parameters to special nodes, regarding their IDs and secret points.

- Provides each node a different pseudo ID, $PSID_i$, and their corresponding secret point $PSSP_i$, which is defined as $PSSP_i = \varphi H1(PSID_i)$; if $i \neq j$ then $PSID_i \neq PSID_j$ as well as $PSSP_i \neq PSSP_j$.

4.2 Anonymous Intra MANET Communication

In anonymous communications, the message content should not consist of any explicit information such as the message sender and recipient addresses. Everything is embedded into the anonymizing message payload. The administrator selects a set of security parameters for the entire system, before the network deployment, including a large prime p and a generator α of Z_p^* . The network nodes A_i , $1 \leq i \leq n$, the corresponding public keys y_i , $1 \leq i \leq n$ of the n participating nodes, x_i 's are randomly selected private keys of A_i 's, where $x_i \in Z_p$, $1 \leq i \leq n$, then y_i is computed from $y_i = \alpha^{x_i} \text{mod} p$ [5]. In this paper we adopted the same anonymous local communication for our network and also the dynamic local MANET formation including the node joining and leaving process from [5].

4.3 Anonymous Communications between Two Arbitrary Special Nodes

When Anonymous authentication is present, two nodes in the same group can authenticate each other *secretly* in such a way that each party reveals its group membership to the other party if and only if the other party is also a group member.

The scheme consists of a set of special nodes and an administrator who creates groups and enrolls special nodes in groups. For this purpose, the administrator will assign each special node A , a pseudonym $PSID_A$ and their corresponding secret point $PSSP_A$, which is defined as $PSSP_i = \varphi H1(PSID_i)$; if $i \neq j$ then $PSID_i \neq PSID_j$ as well as $PSSP_i \neq PSSP_j$. For a given set of $\langle PSID_A, PSSP_A \rangle$, no one can determine the system secret φ . When the special node A wants to authenticate anonymously to the special node B , the following secret hand shake can be conducted. This means that two special nodes A and B can know each other's group membership only if they belong to the same group. When the special node A wants to authenticate to the special node B , the following secret handshake can be conducted which is shown in Figure 1. Pseudo IDs of the nodes are generated considering Pairing-based Cryptography based on bilinear mapping [12].

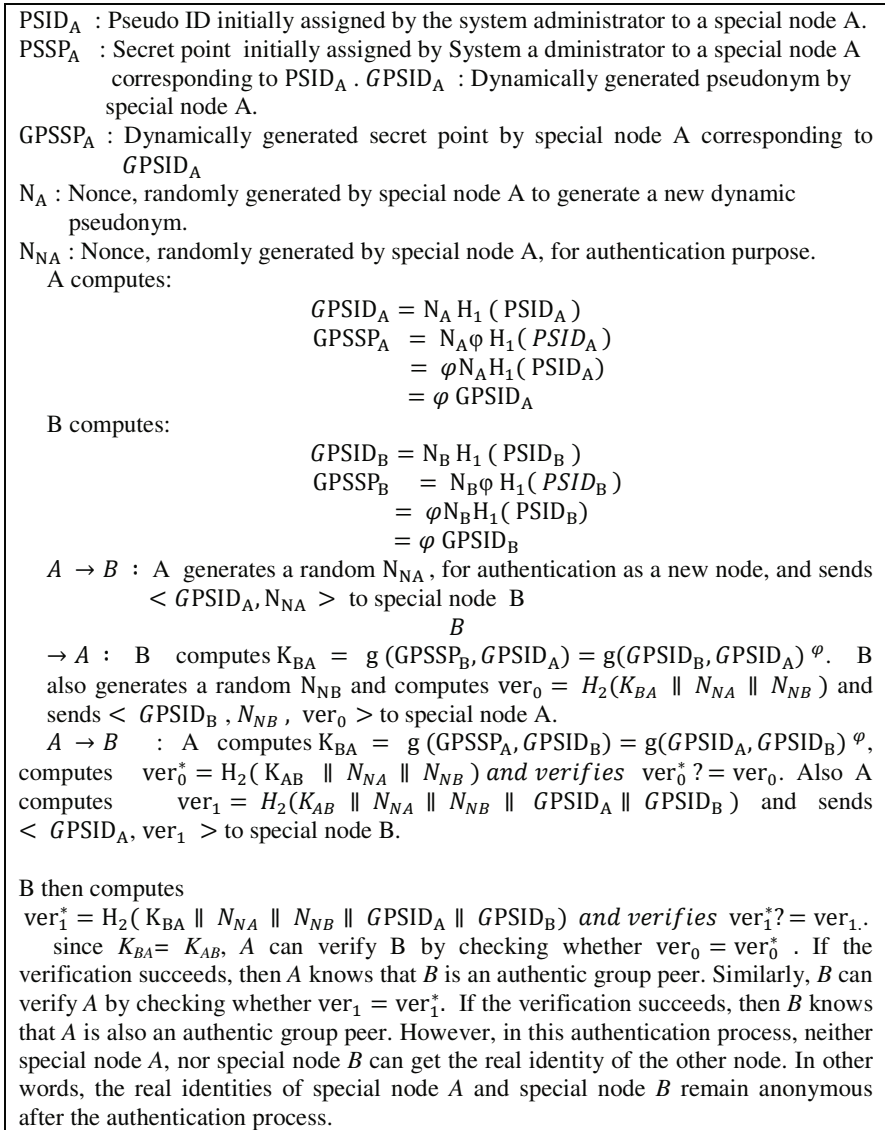


Fig. 1. Anonymous authentication process between two arbitrary special nodes A and B

4.4 Anonymous Communication between Two Arbitrary Ordinary nodes

The sender first randomly selects a local special node and transmits the message as discussed above. On receiving the message, the local special node first determines the destination MANET ID by checking the message recipient flag Fr. If it is 0, then the recipient and the special node are in the same MANET. Otherwise they are in a different MANET. The communication is done using the procedure discussed above.

While providing message recipient anonymity, the message can be encrypted to achieve confidentiality. The presented anonymous communication is quite general and can be used in a variety of situations for communication anonymity in MANET including anonymous file sharing.

5 Security Analysis

We study several attacks designed [5] to analyze the security of the privacy preserving communication protocol.

5.1 Anonymity

(I) It is computationally infeasible for an adversary to identify the message sender and recipient on the local MANET. Hence the privacy preserving communication protocol provides to the sender and recipient anonymity in the local MANET.

(II) The presented communication protocol offers both message sender and recipient anonymity among any two special nodes.

As told earlier, each special node is being assigned a large set of pseudonyms. A dynamically selected pseudonym will be used for any two ordinary nodes in different MANETs to communicate anonymously. The pseudonyms do not carry user information implicitly. The communication can be broken into three segments:

- the communication between the sender and local special node in the message sender's local MANET
- the communication between the special nodes in the corresponding MANETs
- the communication between the recipient special node and the receiver.

(I) has assured the communication anonymity between a special node and an ordinary node in the local MANETs. Therefore we only need to ensure anonymity between two special nodes in different MANETs in order to achieve full anonymity between the sender and receiver.

5.2 Impersonation Attacks

As told above, the forgery attack performed by an adversary, to carry out an impersonation attack is infeasible. For an adversary to forge as a special node, he needs to authenticate himself with a special node A. For this the adversary \mathcal{A} needs to compute $g(GPSID_{\mathcal{A}}, GPSID_B)^\varphi$, where $GPSID_{\mathcal{A}}$ is the identity of the adversary and $GPSID_B$ is the newly generated pseudonym of the special node B.

However, since the adversary does not know the secret point φ , he is unable to compute $g(GPSID_{\mathcal{A}}, GPSID_B)^\varphi$ and impersonate as a special node.

5.3 Message Replay Attacks

Each message packet in communication has a unique one-time session ID (nonce) to protect it from being modified or replayed. In addition, these fields are encrypted

using the intermediate receiver node's public key so that only the designated receiver nodes can decrypt the message. In this fashion, each packet transmitted across different MANETs bears different and uncorrelated IDs and content for adversaries.

Even if the same message is transmitted multiple times, the adversary still cannot link them together without knowing all the private keys of the intermediate nodes.

6 Performance Analysis

In this section we will provide comparison results based on energy consumed by the network nodes of our proposed communication protocol and the method discussed in [7].

Using the method discussed in [7], energy consumed by normal nodes and super nodes are as follows:

At normal node,

$$E_N = E_{SAMAS} + E_{TPKT} + E_D + E_e$$

At super node,

$$E_S = E_{SAMAS} + E_{TPKT} + E_D + E_e + E_{PM} \quad (\text{intra group communication})$$

$$= E_{SAMAS} + E_{TPKT} + E_D + E_e + E_{PM} + E_{Auth} \quad (\text{inter group communication})$$

Where,

E_N : Energy consumed by normal node

E_S : Energy consumed by super node

E_{SAMAS} : Energy required for a node to generate SAMAS message

E_{TPKT} : Energy required transmitting a message

E_D : Energy required for decrypting the message

E_e : Energy required for encrypting the message

E_{Auth} : Energy required by a node to perform authentication process

E_{PM} : Energy required by a super node to maintain set of pseudonyms assigned by the administrator

Using our proposed method, energy consumed by ordinary nodes and special nodes are as given below:

At ordinary node,

$$E_{ord} = E_{PPMAS} + E_{TPKT} + E_D + E_e$$

At special node,

$$E_{spe} = E_{PPMAS} + E_{TPKT} + E_D + E_e \quad (\text{Intra group communication})$$

$$= E_{PPMAS} + E_{TPKT} + E_D + E_e + E_{PG} + E_{Auth} \quad (\text{Inter group communication})$$

Where,

E_{ord} : Energy consumed by ordinary node

E_{spe} : Energy consumed by special node

E_{PPMAS} : Energy required by a node to generate PPMAS message

E_{PG} : Energy required for a node to generate pseudonyms dynamically

Our proposed method has an advantage that during the intra group communication, energy consumed by special node is same as that of ordinary node. During inter group communication also very little energy is required to generate a new pseudonym dynamically. Energy required for maintaining a set of pseudonyms (which is must in [5]) is completely eliminated.

7 Conclusion

In this paper, we first propose an efficient unconditionally secure privacy preserving message authentication scheme (PPMAS) that can be applied to any number of messages without changing the secret exponents. PPMAS ensures message sender privacy along with message content authenticity. To ensure communication privacy without effecting transmission delay and collusion problems, we then proposed a new and efficient privacy-preserving communication protocol for MANET that can provide both message sender and recipient privacy protection. Security analysis shows that the proposed protocol is secure against various attacks.

References

1. Chaum, D.: Untraceable electronic mail, return addresses, and digital pseudonyms. *Communications of the ACM* 24, 84–88 (1981)
2. Khadir, O.: New Variant of ElGamal Signature Scheme. *J. Contemp. Math. Sciences* 5(34), 1653–1662 (2010)
3. Chaum, D.: The dining Cryptographer's problemml: unconditional sender and recipient untraceability. *J. Cryptology* 1(1), 65–75 (1988)
4. Reiter, M.K., Rubin, A.D.: Crowds: Anonymity for Web Transactions. *ACM Transactions on Information and System Security* 1(1), 66–92 (1998); Symposium on Security & Privacy, Oakland, Calif, USA (May 2003)
5. Ren, J., Li, Y., Li, T.: SPM: Source Privacy for Mobile Ad Hock Networks. *EURASIP J. on Wireless Communications and Networking* 2010, article ID534712, 10 Pages (2010)
6. Reed, M.G., Syverson, P.F., Goldschlag, D.M.: Anonymous Connections and Onion Routing. *J. Selected Areas in Communication Special Issue on Copyright and Privacy Protection* (1998)
7. Danezis, G., Dingedine, R., Mathewson: Mixminion: design of a type III anonymous remailer protocol. In: Proc. of the IEEE Computer Society Symposium on Research in Security and Privacy, Oakland, Calif, USA, pp. 2–15 (May 2003)
8. Möller, B.: Provably secure public-key encryptionfor length-preserving chaumian mixes. In: Joye, M. (ed.) *CT-RSA 2003*. LNCS, vol. 2612, pp. 244–262. Springer, Heidelberg (2003)
9. Pointcheval, D., Stern, J.: Security proofs for signature schemes. In: Maurer, U.M. (ed.) *EUROCRYPT 1996*. LNCS, vol. 1070, pp. 387–398. Springer, Heidelberg (1996)
10. Rivest, R., Shamir, A., Tauman, Y.: How to leak a secret. In: Boyd, C. (ed.) *ASIACRYPT 2001*. LNCS, vol. 2248, pp. 552–565. Springer, Heidelberg (2001)
11. Goldwasser, S., Micali, S., Rivest, R.: A digital signature scheme secure against adaptive chosen-message attacks. *SIAM J. Computing* 17, 281–308 (1988)
12. ElGamal, T.A.: A public-key cryptosystem and a signature scheme based on discrete Logarithms. *IEEE Transactions on Information Theory* 31(4), 469–472 (1985)

Prototype Based Performance Prediction of Web Services

Ch. Ram Mohan Reddy¹, D. Evangelin Geetha², K.G. Srinivasa²,
T.V. Suresh Kumar², and K. Rajani Kanth²

¹ B M S College of Engineering, Bangalore - 19, India

² M S Ramaiah Institute of Technology, Bangalore - 54, India

crams19@yahoo.com, degeetha@msrit.edu, kgsrinivas78@yahoo.com,
tvsureshkumar@msrit.edu, rajanikanth@msrit.edu

Abstract. In recent past every discipline and every industry have their own methods of developing products. It may be software development, mechanics, construction, psychology and so on. These demarcations work fine as long as the requirements are within one discipline. However, if the project extends over several disciplines, interfaces have to be created and coordinated between the methods of these disciplines. Prototype of these applications is developed during analysis phase of Software Development Life Cycle (SDLC). But Methodologies for predicting the performance from UML models is available. Hence, in this paper, a methodology for developing Use Case model and Activity model from user interface is presented. The methodology is illustrated with a case study on Amazon.com.

1 Introduction

System engineering concentrates on the definition and documentation of system requirements in the early development phase. The preparation of a system design, and the verification of the system as to compliance with the requirements, taking the overall problem into account: operation, time, test, creation, cost and planning, training and support, and disposal [12]. Systems engineering integrates all disciplines and describes a structured development process, from the concept to the production to the operation phase and finally to putting the system out of operation. It looks at both technical and economic aspects to develop a system that meets the users' needs. Hence a new combination of technologies is required to cop up with the new demand.

1.1 Importance of UML 2.0

In recent past, many researches and software industry use Unified Modeling Language (UML) for the conceptual and logical modeling of any system because of the advantages it has. UML supports both static and dynamic modeling. UML 2.0 is the newer version, which has more features that can be useful for modeling complex systems also [4],[10].

The standard mechanism that UML provides are adaptable itself to a specific method or model, such as constraints and tagged values. we use UML to design RMA processes because it consider an information systems structural and dynamic properties at the conceptual level more naturally than do classic approaches such as

Entity- Relationship model. This approach for modeling RMA processes yields simple yet powerful extended UML use case and sequence diagrams that represent RMA properties at the conceptual level.

There are three classifications of UML diagram.

Behavior diagram. A type of diagram that depicts behavioral features of a system or business process. This includes activity, state machine, and use case diagrams as well as the four interaction diagrams.

Interaction diagrams. A subject of behavior diagrams which emphasize object interactions. This includes communication, interaction overview, sequence, and timing diagrams.

Structure diagrams. A type of diagram that depicts the elements of specification that is irrespective of time. This includes class, composite structure, and component, deployment object and package diagrams.

1.2 Graphical User Interface.

Over some period graphical user interface have become increasingly dominant, and design of the external or visible system has assumed increasing importance. This has resulted in more attention being devoted to usability aspects of interactive system and a need for development of tools to support the design of the external system. Model and notations are required for describing user tasks and for mapping these tasks on to user interface. The primary purpose of task models is to define the activities of the user in relation to the system, as a means of uncovering functional requirements to be supported by the system. Task model focuses on tasks decomposing and/or task flow. A variety of task description methods have been developed, often involving graphical techniques.

Activity diagram, as supported in UML, falls into the latter category. The user of task models in interface design is limited by a lack of tools and techniques to support their definition and weak linkage to visual interface design.

The emergence of Web services introduces a new paradigm for enabling the exchange of information across the Internet based on open Internet standards and technologies. Using industry standards, Web services encapsulate applications and publish them as services. These services deliver XML-based data on the wire and expose it for use on the Internet, which can be dynamically located, subscribed, and accessed using a wide range of computing platforms, handheld devices, appliances, and so on. Due to the flexibility of using open standards and protocols, it also facilitates Enterprise Application Integration (EAI), business-to-business (B2B) integration, and application-to-application (A2A) communication across the Internet and corporate intranet. In organizations with heterogeneous applications and distributed application architectures, the introduction of Web services standardizes the communication mechanism and enables interoperability of applications based on different programming languages residing on different platforms. Web services are self-describing and modular business applications that expose the business logic as services over the Internet through programmable interfaces and using Internet protocols for the purpose of providing ways to find, subscribe, and invoke those services.

Since early 2006, Amazon Web Services (AWS) has provided companies of all sizes with an infrastructure web services platform in the cloud. With AWS compute power, storage, and other services such as gaining access to a suite of elastic IT infrastructure services can be requested. Using Amazon Web Services, an e-commerce web site can weather unforeseen demand with ease; a pharmaceutical company can “rent” computing power to execute large-scale simulations; a media company can serve unlimited videos, music, and more; and an enterprise can deploy bandwidth-consuming services and training to its mobile workforce [13]. Amazon Web Services delivers a number of benefits for IT organizations and developers alike, including:

- Cost-effective. The uses of AWS have to pay only for what they use, as they use it, with no up-front commitments. As the Amazon Web Services cloud grows, the operations, management and hardware costs shrink.
- Dependable. AWS is a battle-tested; web-scale infrastructure that handles whatever that has been throwing at it. The Amazon Web Services cloud is distributed, secure and resilient, giving reliability and massive scale.
- Flexible. AWS provides flexibility to build any application, using any platform or any programming model. The resources can be controlled and fit into your application.
- Comprehensive. Amazon Web Services gives you a number of services that can incorporate into your applications. From databases to payments, these services help to build any great applications cost effectively and with less up-front investment.

Keeping in view of the above discussion we propose a methodology to transform prototype of Web Services into Use Case model and Activity Model. The remaining part of the paper is organized as follows: In Section 2, related work is presented; Section 3 describes the steps involved in the proposed Methodology; The Methodology is illustrated with the Case Study on Amazon.com in Section 4. In Section 5, the paper is concluded and the future directions of the work are highlighted.

2 Related Work

The evaluation of system specifications early in the software development lifecycle has increasingly gained attention from the software engineering community. Early evaluation of software properties, including non-functional ones, is the important in order to reduce costs in software development before resources have been allocated and decisions have been made. Dependability is one example of an important non-functional property and represents the ability to deliver service that justifiably can be trusted. Discovery of Web Services is of an immense interest and is fundamental area of research in ubiquitous computing. Web Services are Internet-based, distributed modular applications that provide standard interfaces and communication protocols aiming at efficient and effective service integration. A Web Service is defined as a functionality that can be programmatically accessible via the Web [11].

A fundamental objective of Web Services is to enable the interoperability among different software applications that run on a variety of platforms [9], [14]. The interoperation has been enabled by the tremendous standardization effort to describe, advertise, discover and invoke Web Services [3]. Web Services are increasingly being adopted as a framework to access Web-based applications. Most of the proposed composition languages for Web Services are based on XML [1], and although XML-based representations have their advantages as universal representations and exchange formats, they can be difficult to understand and to write for non-XML experts. Thus, the use of a graphical modeling language can be very useful to understand the behavior of Systems.

The motivation in this [8] is the use of the Unified Modeling Language (UML) [7], and more specifically the UML Profile for modeling Real Time Systems (RT-UML) profile, as a graphical modeling language for XML Real-Timed Web Services composition and the verification of these systems by using Model Checking techniques on Timed Automata. Some web services workflow patterns and extension in UML to model these patterns are provided in [5]. The patterns identified are Web Service Call, Loop, Data Transformation and Alternate Services. The proposed solutions contain UML activity diagrams along with required extensions. The paper addresses the service composition patterns, but does not provide any feedback mechanism for refinement. [2] talks about MDA approach for development of Web Services and their compositions. Business processes are shown by activity diagrams and static structures by UML class diagrams. Authors have shown mapping from UML to BPEL4WS, WSDL and Java Platform. The UML is a graphical language for visualizing, specifying, constructing, and documenting the artifacts of a software-intensive system [6]. The UML is rather software-specific and strongly characterized by object orientation. While modeling in systems engineering is interdisciplinary. The use of UML can easily lead to acceptance problems and misunderstandings when in interdisciplinary communication. This led to develop our methodology. All these works are referred to developing UI for software system but not for system engineering. Hence based on [4] we have proposed our methodology.

3 Methodology

- i. Consider the prototype of a web service.
- ii. Identify User Interface (UI) elements in the selected prototype
- iii. Develop the flow diagram of the UI elements
- iv. Develop Activity Model
- v. Develop Use Case Model
- vi. Refine & Iterate.

4 Case Study on Amazon.com

4.1 Prototype of Amazon.com Web Service

As an illustration of this methodology, consider the web site Amazon.com, and apply the algorithm for the module Login. The prototype of the login page is given in Fig. 1.

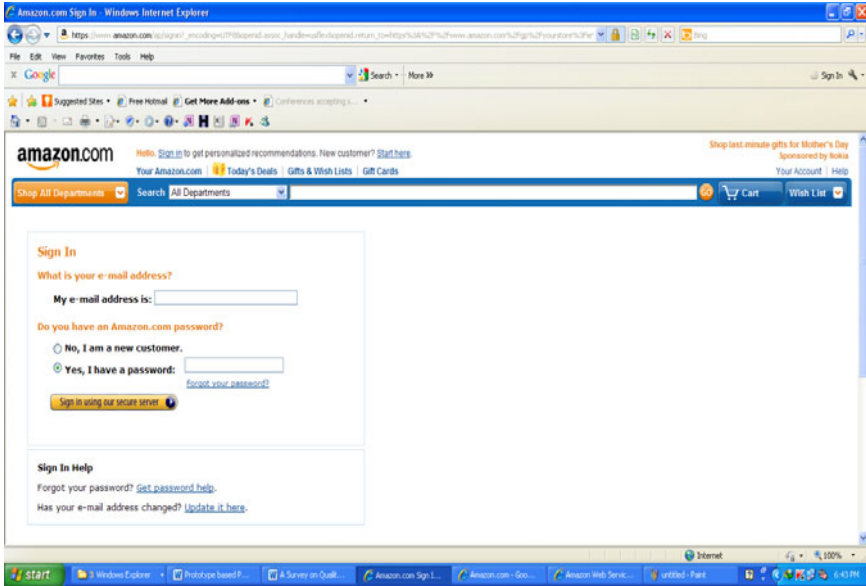


Fig. 1. Screen shot of login page of Amazon.com

4.2 UI Elements of Login

In Fig. 2, the UI elements column identifies four workspaces (W1- W4) as being required to support the user and system activities shown in the three columns. It also identifies two functional elements - Sign in using secure server, Create Account, which are required to support the user tasks. There are two sub-flows in the figure 2, and one exception flow is shown following invalid user input. The information to be displayed is shown in each workspace. This is derived from Prototype model.

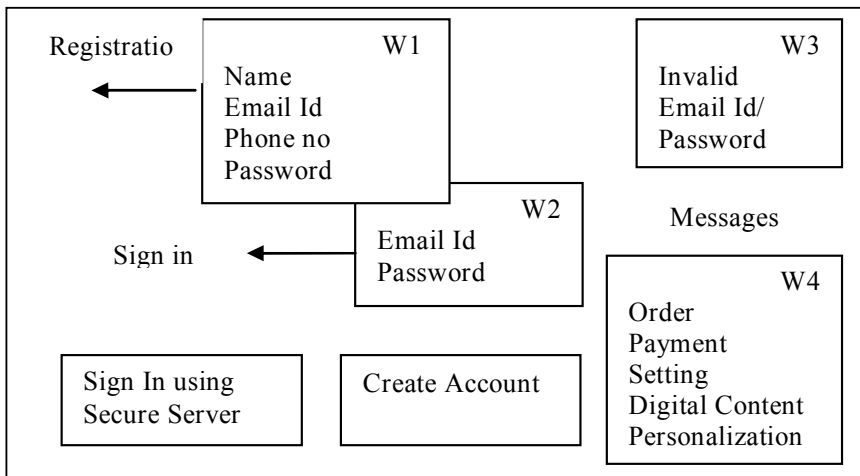


Fig. 2. UI Element cluster for the Login

4.3 Flow Diagram of the UI Elements

4.3.1 Main Flow

In the given below figure 3 showing the ‘Main flow of login’, where the first workspace W1 showing to create a account by providing the Name, Email id, Phone number and password. When it successfully created then workspace W2 showing how to login, for login we have to provide the Email-id and password. If Email-id and password is correct then it enters in secure server. The last workspace W4 shows entering in to system and can access the functionalities.

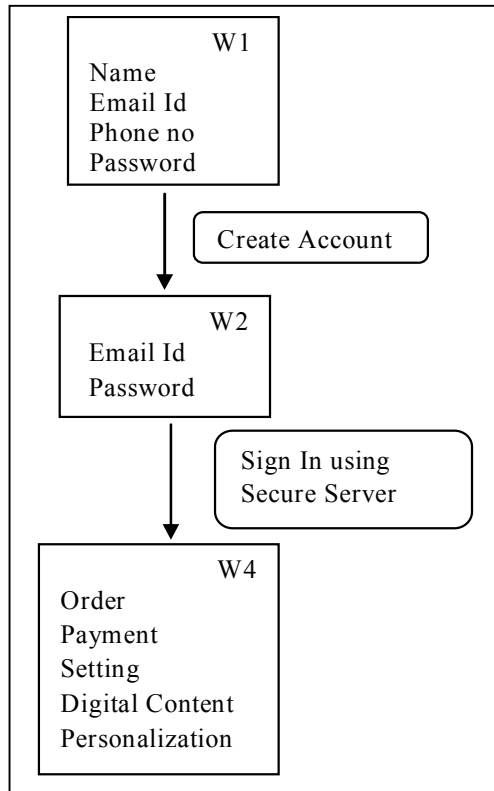


Fig. 3. Main flow of login

4.3.2 Exception Flows

In Fig. 4, the workspace W3 is showing Exception handling when user is logging in. If Email-id or password or both are Incorrect, then it will show the Exception.

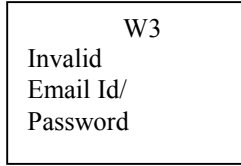


Fig. 4. Exception flows of login

4.4 Activity Model

In Fig. 5, it is showing the activity model for login. When user wants to enter in to the system, if he is new user then he has to provide the Email-id and fill all the details, then he will enter into the system. If he is the existing user then he has to provide email-id or password. If email-id or password or both is incorrect then display error message and again back to login page. If e-mail-id and password is correct then it moves on to open user form.

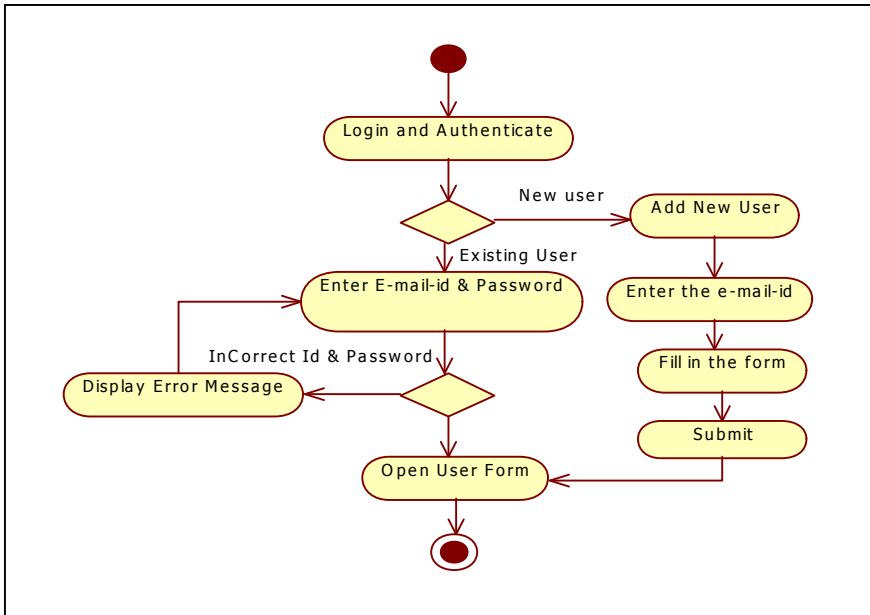


Fig. 5. Activity model for login

4.5 Use Case Model

In Table 1, it is showing use case specification for login and authenticate. It is showing all the description about the use case as who is actor, pre-conditions, post-conditions, basic steps of login, exception handling.

Table 1. Use Case Specification – Login and Authenticate

Brief Description	This use case describes the process by which a user logs into the System.
Actor	User
Pre-Conditions	<p>Precondition One The User must have a valid user name and password.</p> <p>Precondition Two The User has access to the system.</p>
Post-Conditions	<p>Post-Condition One The User successfully logs into the system and is re-directed to the page.</p> <p>Post-Condition Two All and/or any of the login credentials of the User are not valid. The user is informed that the login are invalid and to try again.</p>
Basic Steps	<ol style="list-style-type: none"> 1. If new user, the user should enter the credentials and the system redirects the user to the web page. 2. Already existing user, The user enters and submits his/her Username and password. 3. The System validates and authenticates the user Information. 4. The System re-directs the user to the web page.
Exception Flows	<ul style="list-style-type: none"> • The system displays a message to the user indicating that the user name and/or password are incorrect and to try again. • The system displays a message to the user indicating to enter user name and password. • The user continues at basic step #2. • The system displays a message to the user indicating a communication error. • The user exits the system.

We have obtained the remaining UI Elements and Activity models for Search, shopping cart, online payment and Wish list respectively of amazon.com using the similar procedure.

5 Conclusion and Future Work

In general, developing prototype for software applications is industry practice. Various methodologies facilitate to predict performance from UML models. Hence, in this paper, we have proposed a methodology to transform a prototype into UML models, Activity diagram and Use Case diagram. As future work, we propose to develop methodologies to analyze performance of Web Services from these UML models.

References

1. BPEL 1.1, <http://ibm.com/developerworks/library/ws-bpel>
2. Bézivin, J., Hammoudi, S., Lopes, D., Jouault, F.: Applying MDA Approach for Web Service Platform. In: The 8th IEEE International Enterprise Distributed Object Computing Conference, EDOC 2004 (2004)
3. Curbera, F., Duftler, M., Khalaf, R., Nagy, W., Mukhi, N., Weerawarana, S.: Unraveling the web services web: An introduction to soap, wsdl, and uddi. *IEEE Internet Computing* 6(2), 86–93 (2002)
4. Phillips, C., Kemp, E., Kek, S.M.: Extending UML Use Case Modeling to Support Graphical User Interface Design. In: Proceedings of the 13th Australian Conference on Software Engineering, pp. 48–57. IEEE Computer Society, Los Alamitos (2001)
5. Grønmo, R., Solheim, I.: Towards Modeling Web Service Composition in UML. Presented at The 2nd International Workshop on Web Services: Modeling, Architecture and Infrastructure (WSMAI-2004), Porto, Portugal (2004)
6. Booch, G., Rumbaugh, J., Jacobson, I.: *The Unified Modeling Language, User Guide*. Addison Wesley, Reading (1999)
7. Lara, R., Roman, D., Polleres, A., Fensel, D.: A Conceptual Comparison of WSMO and OWL-S. In: Zhang, L.-J., Jeckle, M. (eds.) *ECOWS 2004*. LNCS, vol. 3250, pp. 254–269. Springer, Heidelberg (2004)
8. Emilia Cambroner, M., Jose Pardo, J., Diaz, G., Valero, V.: Using RT-UML for Modeling Web Services. In: *SAC 2007*, March 11-15. ACM, Seoul (2007)
9. Medjahed, B., Benatallah, B., Bouguettaya, A., Ngu, A.H.H., Elmagarmid, A.: Business-to-Business Interactions: Issues and Enabling Technologies. *The VLDB Journal* (2003)
10. White, S., Cantor, M., Friedenthal, S., Kobryn, C., Purves, B.: Panel: Extending UML from Software to Systems Engineering. In: 10th IEEE International Conference and Workshop on the Engineering of Computer-Based Systems (ECBS 2003), April 07-10, pp. 271–273 (2003)
11. Tsur, S., Abiteboul, S., Agrawal, R., Dayal, U., Klein, J., Weikum, G.: Are Web Services the Next Revolution in e-Commerce (Panel). In: *Proc. VLDB*, pp. 614–617 (2001)
12. Weilkiens, T.: *Systems Engineering with SysML/UML Modeling, Analysis, Design*. In: MK OMG. Elsevier, Amsterdam (2006)
13. Web URL - <http://aws.amazon.com/what-is-aws/> (accessed on 2011)
14. Vinoski, S.: Web Services Interaction Models, Part 1: Current Practice. *IEEE Internet Computing* 6(3), 89–91 (2002)

The Goal of Securing Mobile Ad-Hoc Network and Solutions

Najiya Sultana¹ and S.S. Sarangdevat²

¹ Research Scholar and ² Director, Computer Science & IT Department,
Rajasthan Vidhyapith University, Udaipur, India
saara.sultan@gmail.com, drsssarangdevat@yahoo.com

Abstract. In this article we assess security threats and summarize representative proposals in the context of ad-hoc networks. In this paper, we survey the state-of-the-art approach to providing security for wireless networking paradigms, namely mobile ad-hoc networks. We identify the security threats as well as examine the current solutions. We further summarize lessons learned, discuss open issues, and identify future directions.

1 Introduction

Due to its self-organized nature, the ad-hoc network presents additional security challenges. Because a malicious entity can readily function as a router that forwards packets for peer nodes, it can attack data delivery service through both the control plane (e.g., ad-hoc routing) and the data plane (e.g., multi-hop forwarding). In recent years, wireless networking has been experiencing an explosive growth which resembles the rapid growth of the Internet itself in mid 1990's. Wireless networks offer attractive flexibility and coverage to both network operators and users. Ubiquitous network coverage, for both local and wide areas, can be provided without the excessive costs of deploying and maintaining the wires. Mobility support is another salient feature of wireless networks, which grants the users not only "anytime, anywhere" network access but also the freedom of roaming while networking. Recent advances in wireless communication technology have offered ever increasing data rates, in some cases comparable to their wired counterparts. There are multiple forms of wireless networks with different characteristics and application domains. In this article we focus on wireless networking paradigm, namely ad-hoc networks. Several 3G networks became operational since last 2003. In the absence of a deployed infrastructure, multiple wireless devices can self-organize into an ad-hoc network and forward data through multihop wireless links to enable peer-to-peer communication or further extend the network coverage.

2 Security Goals

To protect the data delivery functionality of network, we focus on the following two security goals:

- *Information Security*, i.e., to provide confidentiality, integrity, authentication, and non-repudiation for two entities that communicate with each other.
- *Network Security*, i.e., to protect the networking system as a whole and sustain its capability to provide connectivity between communicating entities.

Although most, if not all, security threats against the TCP/IP stack in a wired network are equally applicable to an IP-based wireless network, the latter possesses a number of unique vulnerabilities which make it more challenging to secure:

- *Open Wireless Access Medium*: The security threats of message eavesdropping and injection are universe in any network; however they are particularly severe in wireless networks due to the open wireless medium. With off-the shelf hard ware's and little efforts, an attacker can intercept and inject traffic through a wireless channel. There is no physical barrier to separate the attacker from the network, as is the case in wired networks.

Limited Bandwidth: Wireless networks are particularly vulnerable to Denial-of-Service (DoS) attacks due to their limited bandwidth and in-band signaling. Although the wireless channel capacity ¶is continually increasing, the spatial contention problem poses a fundamental limit on network capacity. One can deploy redundant fibers, but everyone must share the same wireless spectrum.

- *System Complexity*: Generally speaking wireless networks are far more complex than their wired counterparts due to the special needs for mobility support and efficient channel utilization. Each piece of complexity added into the system can incur potential security vulnerability, especially in a system with a large user population and a complex infrastructure, such as 3G networks.

In addition, different from wired networks where each host must obtain a *physical* connectivity to the system, an ad-hoc network can be free to join and leave, and the participating entities may not have any *a priori* trust relationship. Due to the open wireless medium, wireless networks require explicit mechanisms to control the membership; the absence of a stringent admission control mechanism may open the door for potential security threats.

After surveying the security threats and proposed solutions in ad-hoc networks, we summarize the lessons learned and shed light on future research directions.

3 Security for Mobile Ad-Hoc Networks

In this section we describe the security issues in a mobile ad-hoc network. We first present the network architecture and routing protocols, then discuss their unique vulnerabilities, followed by an overview of existing solutions for control and data-plane security, as well as key management, in such networks.

Ad-hoc Network Architecture and Routing: Mobile ad-hoc networks (MANETs) enables wireless communications among mobile nodes when a network infrastructure, such as the WLAN or 3G, is unavailable. In a MANET, any node may function as both a data source and a router that forwards packets for other nodes. Packets from a source are typically forwarded through multiple wireless hops to reach the destination. Thus an ad-hoc routing protocol is needed to build and maintain a routing table at each node.

There are mainly three families of routing protocols, namely distance vector, link state, and source routing. In distance vector routing, local neighbors exchange their distance, measured by a given routing metric (e.g., hop count), to each destination. A node selects the neighbor with the shortest distance as its next hop to forward data traffic. In link state routing, each node floods the information about its direct links to the entire network and collects updates from all other nodes to build a complete topology, it then executes a shortest-path algorithm (e.g., Dijkstra) to find a route to the destination. In source routing, each data source specifies the complete path in its data packets; more details can be found in [1].

Each routing approach may have two instantiations: proactive or on-demand. In proactive routing, each node maintains a full routing table for all destinations, and routing updates are exchanged periodically or whenever a topology change occurs. However, this may consume a large portion of limited wireless bandwidth, while only a few routes are actually used. An on-demand approach discovers a route only when needed. The route discovery is initiated by the source, typically through flooding a Route Request (RREQ). Routing states are established at intermediate nodes when RREQ is propagated. Finally the destination sends back a Route Reply (RREP) packet, and data traffic can start to flow.

The three most popular routing protocols for MANETs are DSDV [2] (proactive, distance vector), AODV [3] (on demand, distance vector), and DSR [4] (on-demand, source routing).

4 Security Threats and Goals

Compared with WLANs, the MANETs face a new set of security threats, particularly at the network layer, due to their infrastructure-less nature. Because the core network elements, e.g., routing and forwarding engines, are provisioned by peer nodes, the attacker can readily become a router and disrupt the network operations by attacking the control or data planes. As such, the security vulnerabilities of MANETs present not only in each single-hop wireless link, but also the multi-hop forwarding mechanism that glues these links together as a network.

We focus on the unique network layer security threats in MANETs, which can be divided into two main categories:

Control Plane Attacks: An attacker can announce false routing messages to disrupt the discovery and maintenance of routes between two nodes multi-hop away from each other. The implementation is specific to the ad-hoc routing protocol in use.

Data Plane Attacks: Regardless of which routing protocol is in use, an attacker can drop the data packets passing through it, replay the previously recorded packets, or inject forged packets into the network.

To protect its basic functionality of delivering packets from one node to another, the MANET needs to secure both the control-plane routing and the data-plane forwarding operations.

Accordingly, a complete solution to achieve this goal has at least three pieces: control-plane security, data-plane security, and supporting components such as key management.

5 Control Plane: Secure Ad-Hoc Routing

The control-plane security design typically secure routing protocols by associating various authentication techniques with critical fields (e.g., hop count, source route) in the routing message, thus prevent even both insider and outsider nodes from disrupting the routing functionality.

Authentication Primitives: There are three popular message authentication primitives. The first one is *Message Authentication Codes (HMAC)*. If two nodes share a secret key, they can generate and verify a HMAC for any message using an efficient one-way hash function. However, the HMAC can be verified only by an intended receiver, thus unappealing for authenticating broadcast messages. In addition, it is non-trivial to establish network-wide pair wise keys, as $n(n-1)/2$ keys have to be maintained in a network with n nodes.

One-way key chain is another way to use such one-way functions to authenticate messages. By applying the one-way function $f(\cdot)$ repeatedly on an initial input x , one can obtain a chain of keys

$f^{-1}(x)$. The sender can gradually reveal the key chain in the *reverse* order, and use an unreleased key to generate HMAC for its message. This way, when the key is revealed later, the receiver can use it to verify the HMAC. With time synchronization and careful key release schedule, the one-way key chain can be used to efficiently authenticate broadcast messages [5].

The third authentication primitive is *digital signature* which is based on public-key cryptography, e.g., RSA. Because the message is signed by the sender's secret key, it can be verified by any node given that the public keys have been distributed. Thus digital signature scales well to large numbers of receivers, as in the broadcast cases. However, it involves heavy computation, which presents a new DoS vulnerability. The key revocation is also difficult in a self-organized MANET.

Distance vector routing: To secure distance vector routing protocols such as DSDV and AODV, the main challenge is to ensure that each node advertises the routing metric correctly. For example, when hop count is used as the metric, each intermediate node should increase it exactly by one when the routing updates are propagated in the network.

The hop count hash chain [6], [7] scheme can prevent an intermediate node from *decreasing* the hop count. Assuming the maximum hop count (i.e., the network diameter) is n , a node generates a hash chain of length n each time it initiates a routing update (in DSDV) or a Route Reply (RREP) message (in AODV):

$$h_0, h_1, h_2, \dots, h_n$$

Where $h_i = H(h_{i-1})$ and $H(\cdot)$ is a well - known one-way hash function. The node then adds $h_x = h_0$ and h_n into the routing message, with HOP-count set to 0 and broadcasts it.

When an intermediate node receives such an update, it first check if

$$h_n = H^{n-Hop-count}(h_x)$$

Where $H^m(\text{ho})$ denotes the result of applying $H(\cdot)$ m times on h_x . Then the node updates h_x with $H(h_x)$, increments the HOP count by 1, and re-broadcasts the routing update.

This way, the attacker can never decrease the hop count in an routing update, which provides authentication for the lower bound of the hop count. However, it does not prevent an attacker from advertising the *same* hop count as the one it has received. In [8], a more complicated mechanism called hash tree chain is proposed to ensure a monotonically increasing hop count as the routing update traverses the network. One general limitation of the above approaches is that they are only applicable for *discrete* routing metrics, while ineffective for continual metrics that take non-integer values.

Note that non-cryptographic techniques can also be used to secure distance vector routing protocols. For example, in the context of RIP [9], the proposed solution in [10] exploits simple geometry of triangle theorem as follows: For any set of three nodes A , B , and C , the distance between A and B should be no more than the sum of the distance between A and C , and the distance between B and C . A node can use the above fact to detect the inconsistency among the received routing updates, and further use TTL-controlled ICMP probe messages to verify the suspicious distances.

Link state routing: To secure link state routing protocol such as OSPF, the main challenge is to prevent the forgery of non-existent links. This is achieved in Secure Link State Routing (SLSP) [11] through digital signature based authentication. In SLSP, each node periodically broadcasts *hello* messages for local neighbor discovery, and floods *Link State Update* (LSU) packets to advertise its links. Both *hello* messages and LSU packets are signed by the node's private key. A link is accepted into the global topology if and only if it is advertised by both end-nodes through valid LSU packets. Thus the attacker cannot forge any link that involves a legitimate node.

The DoS vulnerability that exploits the high computation overhead of digital signature is addressed through rate control mechanisms. Each node measures how frequently its neighbor sends the digitally signed control packets, and discards such packets without verification if the rate exceeds a threshold.

Source routing: To secure source routing protocols such as DSR, the main challenge is to prevent malicious manipulation of the source routes (i.e., an ordered list of intermediate nodes) by an intermediate node, e.g., addition of new nodes, removal of existing nodes, or order switching. This is typically achieved via a per-hop authenticator associated with the route.

Ariadne [12] is a secure extension of DSR. It uses one-way HMAC chain, i.e., TESLA [5], to authenticate the source routes. It assumes time synchronization and pre-distribution of the last key in each node's TESLA key chain. Take the following example as an illustration: the source node S has a route towards the destination D through three intermediate nodes A , B , and C . When the Route Request (RREQ) packet is propagated, each intermediate node appends itself to the source route, together with a hash value for the entire packet, and a HMAC keyed by its next unreleased TESLA key. When the destination receives RREQ, it verifies whether the content matches the hash value. If so, it appends the cascaded HMACs in a Route Reply (RREP) packet, which traverses the reverse path to the source. An intermediate node delays the RREP packet until it released the previously used TESLA key, so that the next hop node can verify its TESLA HMAC.

The per-hop hash prevents a malicious node from modifying the RREP packet that it has received. Thus all it can do is to append new nodes to the route. The per-hop TESLA HMAC further prevents it from adding any nodes other than itself, because it does not know the unreleased keys of other nodes. This way, the source route discovered by Ariadne is secure.

6 Data Plane: Protect Packet Forwarding

The data-plane security should ensure that each node forwards packets according to its routing table. Unlike the control plane, the data plane cannot be proactively secured by cryptographic primitives because several attacks on forwarding cannot be prevented: an attacker may simply drop all packets passing through it, no matter how well they are protected. Thus the security solution takes a reactive approach, at the heart of which is a detection technique and a reaction scheme.

Detection: The open wireless medium enables localized detection in MANETs, in which each node overhears the channel and monitors the behavior of its neighbors. However, its accuracy is limited by a number of factors such as channel error, mobility, hidden terminals, etc. A malicious node may even abuse the detection mechanism and intentionally accuse legitimate ones. To address such issues, the detection results at individual nodes can be synthesized in a distributed manner to achieve consensus among a group of nodes. An alternative detection approach relies on explicit acknowledgment from the destination, and/or intermediate nodes, to the source, so that the source can figure out where the packet was dropped.

Localized detection: The *watchdog* technique [13] takes the localized approach to detecting misbehaviors in the context of DSR. It assumes symmetric links, i.e., if A can hear B , then B can also hear A . Since the entire path is specified, when node A forwards a packet to the next hop B , it knows B 's next hop C . It then overhears the channel for B 's transmission to C . If it does not hear the transmission after a timeout, a failure tally associated with B is increased. If the tally exceeds a threshold bandwidth, A reports B 's misbehavior to the source.

This concept is extended in [50] to work with distance vector protocols such as AODV. It adds a *next hop* field in AODV packets so that a node can be aware of the correct next hop of its neighbors. Independent detection results are authenticated and further synthesized to reach consensus among local neighbors. It also considers more types of attacks, such as packet modification, duplication, and jamming attacks.

ACK-based detection: The detection mechanism in [15] is based on explicit acknowledgments. The destination acknowledges each received packet. Based on the delivery quality, the source can initiate a fault detection process on a suspicious path. It performs a binary search between

itself and the destination, and sends out data packets piggybacked with a list of intermediate nodes, also called "*probes*", which should send back acknowledgments. The source shares a key with each probe and the probe list is "onion" encrypted. Upon receiving the packet, each probe sends back an ACK, encrypted with the key shared with the source. The source in turn verifies the encrypted ACKs and attributes the fault to the node closest to the destination that sends back an ACK.

Reaction: Once a malicious node is detected, the network should be protected by the reaction scheme to prevent future attacks from it. The reaction scheme is typically related to the prevention component in the overall security system. For example, the malicious node may be revoked of its certificate, or have lower chance to be chosen in future forwarding paths. Based on their scope, the reaction schemes can be categorized as network-wide reaction and end-host reaction.

Network-wide reaction: The network-wide reaction in [14] is based on the URSA certification framework [16]. Once multiple nodes have independently detected that one of their neighbors is malicious, they collectively revoke its current certificate. Consequently, the malicious node is isolated in the network as it cannot participate in neither routing nor packet forwarding operations any more.

End-host reaction: The *path rater* in [13] is an end-host reaction scheme that allows each node to maintain its own rating for another node. A node slowly increases the rating of well-behaving nodes over time, but dramatically decrease the rating of a malicious node that is detected by its *watchdog*. Based on the rating, the source always picks up a path with highest average rating. Clearly each node may have different opinion about whether another node is malicious, and make its independent reaction accordingly.

7 Supporting Components: Trust and Key Management

There are two supporting components indispensable in any network security solutions, namely trust and key management. The basic functions provided by them include trust establishment, key generation, key exchange, and key update. The simplest solution relies on *a priori* trust relationship and pre-distribution of keys among the nodes. However, this may not apply to a self-organized MANET in which the nodes do not have any *a priori* bindings. Instead, such networks require in-the-band solutions that can establish trust and keys on-the-fly. The trust and key management proposals for MANETs are mainly concerned with a hybrid of asymmetric and symmetric cryptosystems, where trust is established via public-key credentials or tickets, and shared symmetric keys are exchanged for later use. An inherent issue in trust management is a *trust graph*, where the nodes correspond to the network entities and edges to the verifiable credentials. A directed edge is added from node X to node Y if X 's credentials are verifiable by Y . The concept of trust graph provides a means for classifying the various key management solutions into three categories, as we elaborate next.

Trusted Third Party (TTP): In this paradigm, a centralized authority (e.g., a KDC or CA) is trusted by all entities. An entity A is trusted by another if the authority claims that A is trustworthy. This scheme is centrally managed and vulnerable to single point of failure. In addition, the neighborhood of the central point is potentially the bottleneck of a scalable network and subject to denial-of-service attacks.

The certification service proposed in [17] distributes among n special nodes designated as the *server infrastructure*. It does not use a single centralized server, nor a server hierarchy as is the current Internet practice [18]. Instead, it employs threshold cryptography [19] to distribute n shares of the private key of certification service among the server nodes, while the respective public key is assumed to be well -

known. Consequently, the ability to sign certificates is also shared, and performed by generating and combining partial signatures from $t+1$ servers. The system is able to tolerate up to t server compromises. In order to protect against mobile adversaries [20] and adapt its configuration to network dynamics, *proactive share refreshing* periodically updates the shares without disclosing the service private key to any subset of less than $t+1$ servers.

“Web-of-trust”: In this paradigm, no particular structure or commonly trusted entity exists in the trust-graph [21]. Instead, each entity manages its own trust based on direct relationship and recommendation from others. The scheme is fully distributed, making it resilient to attacks, but also difficult to achieve consensus among various entities.

A distributed, self-organized public-key management system that follows this concept is proposed in [22][23], in which every node acts as its own authority. However, there is no well-known certificate directory where all certificates are stored, as opposed to PGP [21]. Instead, each node maintains a small local certificate repository. When two nodes need to verify the public keys of each other, they merge their local repositories and apply a shortcut search algorithm to find a trust chain.

The bootstrapping of “web-of-trust” can be built on top of direct communication through secure side channels such as *infras* [24]. Such channels allow two nodes close enough to each other to directly exchange credentials. This idea essentially mimics human behavior, and is similar to that proposed in [25], where a pre-authentication phase between two mobile devices makes possible the secure exchange of public keys of the nodes over location-limited channels. One-hop recommendation on the trust graph is used in [24] to expedite trust establishment. That is, if A and B have established trust with C via direct contact, A and B also trust each other. However, such referral is non-transitive, thus valid only once.

Localized trust: This paradigm [26] is the middle ground between the previous two. A node is trusted if any k trusted entities among its one-hop neighbors claim so, within a bounded time period. As trust management is fully distributed in both space and time domains, it fits well with large-scale, dynamic MANETs with on-demand authentication requirements.

The localized trust paradigm have a particularly desirable feature of localized authentication traffic. Recent analysis [27] shows that the capacity of MANETs rapidly approaches zero as network size increases, in particular when network traffic is *not* localized. This motivates the certification framework in [16] to achieve balance among global consensus on trust, system scalability, and service availability. The framework employs threshold secret sharing to distribute the CA functionality into each and every node. Any $t+1$ local neighbors of a node can collectively sign its certificate, which then can be verified and accepted in the entire network. The localized traffic pattern not only improves the network scalability, but minimizes the impact of multi-hop wireless dynamics on the service availability.

8 Summary

In contrast to WLAN and 3G networks, a MANET does not have a pre-deployed and trusted infrastructure. Therefore, the security solutions are mainly concerned with securing the operations of distributed network protocols, and establishing trust among

peer nodes. Proactive approaches are commonly used to secure the control plane, especially the routing protocols, by authenticating the signaling messages. On the other hand, the data plane is protected through reactive approaches that detect and react to occasional intrusions. The complete security solution should encompass both aspects and manage the required keying materials in a self-organized manner.

9 Conclusion

A. Lessons Learned

To secure ad-hoc networks, however, requires much more complex mechanisms because nodes within the system are considered untrustworthy. Therefore, one must secure both the control plane and the data plane, in addition to ensuring the privacy and integrity of the transmitted data. There are several general observations that can be drawn from this study. First, the cryptographic techniques are an essential ingredient in providing information security, and can serve as the first line of defense against network attacks (e.g., through authentication). However, cryptography alone does not suffice to secure a networking system. Given a specific security requirement, there is neither a systematic process to develop a suitable design nor an automatic way to gauge its vulnerability at this time. One major challenge in network security design is to address various dimensions of performance tradeoff including crypto strength, execution speed, computational overhead, communication cost, energy consumption, operational and configuration complexity, to name a few. A feasible solution must balance among these partially conflicting goals, yet a systematic way to evaluate these tradeoffs is still missing.

Unfortunately, many state-of-the-art solutions still do not possess the properties that networking protocols deem necessary. For example, most of them cannot scale to an unconstrained network scale, and rely on centralized components to bootstrap or operate the security protocols. As a result, they may work well in a small-scale setting but cannot function effectively or efficiently in large-scale networks, such as the emerging sensor networks or metropolitan/community mesh networks.

Finally, the current security solutions are typically based on specific threat models, and operate explicitly or implicitly with a number of assumptions made on the networks. For example 3G network security designs assume a reliable core network. When unexpected failures or unanticipated threats occur, these security solutions are likely to collapse.

B. Future Research

Looking ahead into the future, we would like to identify two directions that need more research and development efforts to build a truly secure wireless networking system:

Critical evaluation: While many security solutions have been reported in the literature, most of them have not been thoroughly evaluated in terms of security strength and system performance. In fact, we lack systematic evaluation methods and efforts in the following aspects: (1) *Vulnerability analysis of the current solutions/standards.* While the crypto strength of individual ciphering algorithms is relatively well understood, we have no formal analytical tools to assess a system security proposal. In particular, the

analysis on the inter-dependency among various system components and security operations poses a major research challenge. (2) *Measurements and emulations*. To date most solutions have been evaluated for their network and system performance via simulations. However, tested measurements, or large-scale emulations based on measurement traces, are critically needed for a more accurate understanding.

Resilient security: Most current solutions make idealistic assumptions on the network and individual components. A truly resilient security solution needs to possess both robustness and resiliency. It must be robust against wireless channel errors, transient/permanent network connectivity and topology changes, and user mobility. It must also be resilient against unanticipated attacks, operational errors such as misconfigurations, and compromised/stolen devices. In addition, both dimensions of goals have to be achieved with acceptable cost.

References

- [1] Kurose, J., Ross, K.: Computer Networking: A Top-Down Approach Featuring the Internet. Addison-Wesley, Reading (2002)
- [2] Perkins, C., Bhagwat, P.: Highly Dynamic Destination-Sequenced Distance-Vector Routing (DSDV) for Mobile Computers. In: Proc. SIGCOMM (1994)
- [3] Perkins, C., Royer, E.: Ad Hoc On-demand Distance Vector Routing. In: Proc. WMCSA (1999)
- [4] Johnson, D., Maltz, D., Jetcheva, J.: Ad Hoc Networking. In: DSR: The Dynamic Source Routing Protocol for Multi-Hop Wireless Ad Hoc Network, ch. 5. Addison-Wesley, Reading (2001)
- [5] Perrig, A., Canetti, R., Tygar, D., Song, D.: The TESLA Broadcast Authentication Protocol. RSA CryptoBytes 5(2), 2–13 (2002)
- [6] Hu, Y., Johnson, D., Perrig, A.: Sead: Secure efficient distance vector routing for mobile wireless ad hoc networks. In: Proc. IEEE Workshop on Mobile Computing Systems and Applications, WMCSA (2002)
- [7] Zapata, M., Asokan, N.: Securing Ad Hoc Routing Protocols. In: Proc. WiSe (2002)
- [8] Hu, Y., Perrig, A., Johnson, D.B.: Packet Leashes: A Defense against Wormhole Attacks in Wireless Networks. In: Proc. INFOCOM (2003)
- [9] Malkin, G.: Routing Information Protocol Version 2. RFC 2543 (1998)
- [10] Pei, D., Massey, D., Zhang, L.: Detection of Invalid Routing Announcements in RIP Protocol. In: IEEE Globecom (2003)
- [11] Papadimitratos, P., Haas, Z.J.: Secure Link State Routing for Mobile Ad Hoc Networks. In: IEEE Workshop on Security and Assurance in Ad Hoc Networks (2003)
- [12] Hu, Y., Perrig, A., Johnson, D.B.: Ariadne: A Secure On-demand Routing Protocol for Ad Hoc Networks. In: Proc. MOBICOM 2002 (2002)
- [13] Marti, S., Giulì, T., Lai, K., Baker, M.: Mitigating Routing Misbehavior in Mobile Ad Hoc Networks. In: Proc. Mobicom (2000)
- [14] Yang, H., Meng, X., Lu, S.: Self-Organized Network Layer Security in Mobile Ad Hoc Networks. In: Proc. WiSe (2002)
- [15] Awerbuch, B., Holmer, D., Nita-Rotaru, C., Rubens, H.: An On-Demand Secure Routing Protocol Resilient to Byzantine Failures. In: Proc. WiSe (2002)

- [16] Luo, H., Kong, J., Zerfos, P., Lu, S., Zhang, L.: URSA: Ubiquitous and Robust Access Control for Mobile Ad-Hoc Networks. *IEEE/ACM Transactions on Networking* (2004) (to appear)
- [17] Zhou, L., Haas, Z.: Securing Ad Hoc Networks. *IEEE Network* 13(6), 24–30 (1999)
- [18] Aresenault, A., Turner, S.: Internet X.509 Public Key Infrastructure. Internet Draft, draft-ietf-pkix-roadmap-06.txt (2000) (work in progress)
- [19] Desmedt, Y.: Threshold cryptography. *European Transactions on Telecommunications* 5(4), 449–457 (1994)
- [20] Ostrovsky, R., Yung, M.: How to Withstand Mobile Virus Attacks. In: *Proc. ACM Symposium on Principles of Distributed Computing* (1991)
- [21] Zimmermann, P.: *The Official PGP User's Guide*. MIT Press, Cambridge (1995)
- [22] Hubaux, J., Buttyan, L., Capkun, S.: The Quest for Security in Mobile Ad Hoc Networks. In: *Proc. MobiHOC* (2001)
- [23] Capkun, S., Buttyan, L., Hubaux, J.: Self-Organized Public-Key Management for Mobile Ad Hoc Networks. *IEEE Transactions on Mobile Computing* 2(1) (January-March 2003)
- [24] Capkun, S., Hubaux, J., Buttyan, L.: Mobility Helps Security In Ad Hoc Networks. In: *Proc. MobiHoc* (2003)
- [25] Balfanz, D., Smetters, D., Stewart, P., Wong, H.: Talking to Strangers: Authentication in Ad-Hoc Wireless Networks. In: *Proc. NDSS* (2002)
- [26] Kong, J., Zerfos, P., Luo, H., Lu, S., Zhang, L.: Providing Robust and Ubiquitous Security Support for Mobile Ad-Hoc Networks. In: *Proc. ICNP* (2001)
- [27] Gupta, P., Kumar, P.: The Capacity of Wireless Networks. *IEEE Transactions on Information Theory*, IT-46(2), 388–404 (2000)

Elements of Cloud Computing: A Perspective on Service Oriented Enterprises (SOEs)

R. Suchithra¹, R. Selvarani², and Dhinaharann Nagamalai³

¹ Master of Computer Applications Department, Jain University

² Computer Science and Engineering Department, MSRIT Bangalore, India

³ Wireilla Net solutions, Victoria, Australia

{suchithra.suriya,selvarani.riic}@gmail.com

Abstract. Cloud computing has made the potential dream of the IT industry come true by providing software as a utility as a better service and transformed the way the hardware has been designed. Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services. It reaps all the benefits of virtualization, grid computing and distributed computing which offers software, infrastructure and platform as a service on the basis of “pay as you use model”. This paper addresses various issues related to the service oriented enterprises in terms of on demand computing (cloud computing). A survey on quality of service provided by various cloud service providers is also presented here with a discussion of various issues in integrating the service oriented enterprises and cloud services.

1 Introduction

Although cloud computing takes some of the old concepts of service-oriented architecture, distributed and grid computing much has been discussed about cloud computing recently. Cloud, which is the next revolutionary step in distributed computing, will change the world in such a way that people can get any service on the cloud where cloud is the limit. The present generation of enterprise customers require geographically and heterogeneously distributed resources such as computing, software, infrastructure or application platform etc. These resources are distributed in multiple locations. Cloud computing enables the creation of virtual organisations by allowing the users easy and transparent access to these resources in a pay-per-use and self-service way. Cloud keeps up the earlier promise of the Internet that you could dynamically move out of another environment. With the abundance of cloud providers in the market who have their own API or set of services, different pricing models, the customers are perplexed as to what to choose [1]. Another challenge is that cloud is a monolithic system where “one provider does everything” and the customer is mostly locked in with a single service. This paper gives a comparative study of various cloud service providers available in the market and the products they offer and their features.

2 Background

Cloud offers basically three types of services: (a) application as a Service, (b) platform as a service and (c) software as a service. There are many service providers like 3Tera, Net suite, Amazon Web services, IBM, Microsoft etc in the market offering cloud services. There is no specific standard where each cloud service provider defines his product in his own way. There are lots of geographical, political as well as security, reliability, payment and performance-based issues in cloud services. The other major issue the user faces is that he is locked with a vendor for a long time. We address these issues by briefly discussing the cloud structure, the major problems and the various cloud service providers in the market today.

3 Framework of Cloud Computing

Many articles and blogs regarding the framework of the cloud address it from the point of the view of the vendors. Here we intend to explain the framework to provide consumers, researchers and academicians a better knowledge of cloud computing, and explain its advantages and the issues in the cloud usage. The taxonomy is briefly explained below:

3.1 Cloud Service Models

Cloud offers basically three types of services: (a) IAAS (infrastructure as a service), (b) PAAS (platform as a service) and (c) SAAS (software as a service). To explain these three S's, let's understand the old web application development model.

i. Infrastructure as a Service

The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications [2]. For example if we want to use a software and we need storage areas, then we have to purchase servers from hardware vendors; if we want to extend the web application for external use, then we want routers, switches, load balancers and we have to buy bandwidth and hosting services. So this infrastructural facility is provided by IAAS vendors. Managed hosting and development environments are the basic services provided by the IAAS vendors. Virtualisation has enabled IaaS providers to offer almost unlimited instances of servers to customers and make cost-effective use of the hosting hardware. It is sometimes referred as Hardware as a service (HAAS). Some of the IAAS vendors are Amazon.com, Google etc.

ii. Platform as a Service

The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider [3]. For example once we have a server we must purchase operating system and we have to build an application server stack and as a next step we need database software. This service is provided by PAAS vendors. PaaS

provides a platform for developers. The end users write their own code and the PaaS provider uploads that code and presents it on the web. Salesforce.com's Force.com is an example of PaaS. PaaS provides services to develop, test, deploy, host and maintain applications in the same integrated development environment. It also provides some level of support for the creation of applications. Thus PaaS offers a faster and more cost effective model for application development and delivery. The PaaS provider manages upgrades, patches and other routine system maintenance. Since PaaS is based on a metering or subscription model, users only pay for what they use. Users take what they need without worrying about the complexity behind the scenes. There are four types of PaaS applications such as (1) social application platforms, (2) raw compute platforms, (3) web application platforms and (4) business application platform. Facebook is a type of social application platform wherein third parties can write new applications that are made available to end users. The CRM solutions provided by the companies are examples of business application platform. Developers can upload and execute their applications on Amazon's infrastructure which is an example of raw compute platform. Google provides APIs to developers to build web applications, which are an example of web application platform.

iii. Software as a Service

It is the service based on the concept of renting software from a service provider rather than buying it yourself. It is known as software on demand. The capability provided to the consumer is to use the provider's applications running on a cloud infrastructure. The applications are accessible from various client devices through a thin client interface such as a web browser [2]. Once the hardware is brought and the operating system is installed, we can install the actual web application and as SaaS applications exist in a broader IT infrastructure, they need to be able to interact with other data and applications. This service is provided by SaaS vendors. The SaaS platform provides flexibility in an application's location -- it can run in a data centre or on the desktop, where so much of the world's computer power lives. This approach enables services to be implemented and consumed where appropriate. This platform's flexibility provides several benefits for IT organisations. Users can benefit from superior application functionality where and when they need -- people are at the centre of this vision. The platform enables scenarios where some applications are delivered as a service, while others run on client and server devices. Yahoo mail, Google docs are examples of SaaS vendors.

There are basically three types of cloud: (1) private and (2) public and (3) hybrid cloud.

1. Public Cloud or External Cloud

It describes cloud computing in the traditional main stream sense, whereby resources are dynamically provisioned on a fine-grained, self-service basis over the Internet, via web applications/web services, from an off-site third-party provider who bills on a fine-grained utility computing basis with the benefits of easy and inexpensive set-up because hardware, application and bandwidth costs are covered by the provider, scalability and no wastage of resources because of the pay-as-you-use model. Compute Cloud (EC2), IBM's Blue Cloud, Sun Cloud, Google AppEngine and Windows Azure Services. Though there is a belief that private clouds dominate over

public clouds among the enterprises in the short run, but in the long run much of the workload will move to the public clouds. The cost elasticity will be a major driving factor for this move. But the public clouds are still not mature because of the various unaddressed security issues.

2. Private Cloud

Private cloud (also called internal cloud or corporate cloud) is a marketing term for a proprietary computing architecture that provides hosted services to a limited number of people behind a firewall [2]. It may be managed by the organisation or a third party, and may exist on premise or off premise, but only members of the client organisation can provision, configure and use the resources of that private cloud. For example, a university could build a cloud that attending students alone would be allowed to use, or a financial services company could build a cloud that only their employees could access. How to build a private cloud depends on what an enterprise is going to start with and the legacy of the environment. We have to virtualise the servers if we are going to start from the dust. As a next step we have to virtualise the storage and network. These are the basic requirements that an enterprise has to strictly follow if it needs to establish a private cloud. The private cloud should offer necessary hardware and software provisions to the customers on their request. The enterprise should develop a method to control and manage the cloud environment and a paradigm has to be developed to move the data between private and public clouds. The enterprise has to decide as to what is going to run in a private cloud and what in a public cloud, and which applications are scalable in a private and which are not.

3. Hybrid Cloud

There are signs that private and public cloud models are giving way to the enterprise hybrid cloud. A hybrid cloud is a cloud computing environment in which an organisation provides and manages some resources in-house and has others provided externally. For example, an organisation might use a public cloud service, such as Amazon's Elastic Compute Cloud (EC2), for general computing but store customer data within its own data centre [10]. Using a hybrid model the enterprise can mix and match the resources between local infrastructures, which is typically a sunk cost but difficult to scale, with infrastructure that is scalable and provisioned on demand. Hybrid cloud will dominate the market in the future than public cloud because most of the enterprises have made huge investments in the infrastructure required to provide resources in-house and they prefer to keep the data in their control and the other services can be given to the external providers [20]. For example in the case of business intelligence, hybrid approach can be followed where keeping the data can be local and the analytical processing in the cloud can be given to the public cloud. Some enterprises have stepped into hybrid cloud. Google recently began offering "applets," downloadable appliances that let users deploy the company's cloud-based offerings internally.

3.2 Virtualisation

Cloud computing is a pool of virtualised resources that has gained the attention and imagination of organisations of all sizes because its service-delivery model converts

the power of virtualisation into measurable business value by adding the provisioning and billing capabilities. Virtualisation is an IT paradigm that separates computing functions and technology implementations from physical hardware. Cloud computing, for example, is the virtualisation of computer programs through an Internet connection rather than installing applications on every office computer [7]. Cloud computing that is a pool of reserved resources allows the resources to be pooled through the provisioning of virtual or physical machines. Best of all, these processes are aimed to be available with very little to no downtime. Virtualisation is efficient because expensive physical devices that were dedicated to specific applications can be shared across a number of functions without modifying the programs. This drives physical asset utilization higher, resulting in more work completed per capital or operating dollar spent, leading to lower operating costs [15]. As enterprises are investing more in virtualised environments, they sometimes find that one of the major hyper visor products fits the requirements of a workload better than the others. As they review their applications, they may find that other workloads might do better in one of the other environments. Moving virtual servers from one environment to another is quite difficult and yet, organisations face this requirement from time to time [17]. But this problem can be pulled off because of a few products like Racemi's Dynacenter available in the market. Though virtualisation is the most disruptive technology in the marketplace which has enabled massive progress in data centre consolidation, storage area networks (SANs), ultra-high speed LANs and cloud computing, virtualisation has resulted in lots of issues that must be tackled. The virtualisation's problem number one is back-up and disaster recovery. While most back-up vendors now come up with good solutions through their products, we may find that it is either a paid add-on or requires an upgrade to their standard package.

3.3 Fear of the Cloud

Transparency, Use Limitation, Disclosure, Security Management System, Customer Security Features, Data Location, Breach Notification, Audit, Data Portability and Accountability are the common insomnia-inducing anxieties for anyone working at enterprise scale. Security concern is a major issue that prevents enterprises from taking advantage of cloud. In this section we discuss taxonomy of various security concerns that a cloud environment faces. A recent survey commissioned by Microsoft has found that 75 percent of senior business leaders consider safety, security and privacy as the top potential risks of cloud computing. More than 90 percent of the general population and business leaders would have concerns about the security and privacy of their own data in a cloud computing environment, according to the survey by Penn Schoen & Berland [12]

i. Data Security

Sensitive data processed outside the enterprise involves a greater level of risk because outsourced services bypass the "physical, logical and personnel controls" when compared with the control measures imposed on the in-house programs. Customers are responsible ultimately to the security of the data even though the data are held by the cloud service provider [16]. Get as much information as you can about the people who manage your data. "Ask providers to supply specific information on the hiring

and oversight of privileged administrators, and the controls over their access,” is what Gartner says. Although customers stress upon external auditing, most of the companies are not doing regular auditing within the organisation to gauge the risk of potential security threats to the business [19]. There is no logging and audit trail information throughout the cloud stack available to cloud customers. It’s necessary to monitor access control logs, virtualisation management logs, hypervisor logs and virtual networking logs to be compliant. Further, if there are logs on the hypervisor system they are likely not segmented by customer, meaning no actionable information can be provided to an auditor. Most of the cloud service providers take enough measure to secure the data by applying data encryption mechanisms, user authentication and authorization practices; but the users are always worried about the vulnerability of external data to hawkers and disgruntled employees. The cloud service providers should ensure that encryption policies are designed and tested by experts.

ii. Data Sanitisation

Sanitisation is one key element in assuring confidentiality. Sanitisation is the process of removing the data from media before reusing the media in an environment that does not provide an acceptable level of protection for the data that have been on the media before sanitising. The major issue a client faces in case of data sanitisation is whether the cloud service provider would implement some sound practices for redundant and retiring data storage devices as and when these devices are retired or taken out of service. Media flows in and out of organisational control through recycle bins in paper form, out to vendors for equipment repairs, and hot swapped into other systems in response to emergencies. This potential vulnerability can be mitigated through proper understanding of where information is located, what that information is and how to protect it. IS resources shall be sanitised before they are released from classified information controls or released for use at a lower classification level. There are different types of sanitisation for each type of media and more discussed methods: disposal, clearing, purging and destroying. Some media can be simply disposed if information disclosure would have no impact on organisational mission, would not result in damage to organisational assets, or in financial loss or harm to any individuals. Clearing information is a level of media sanitisation that would protect the confidentiality of information against a robust keyboard attack. Purging information is a media sanitisation process that protects the confidentiality of information against a laboratory attack. Destruction of media is the ultimate form of sanitisation. After media are destroyed, they cannot be reused as originally intended. Physical destruction can be accomplished using a variety of methods, including disintegration, incineration, pulverising, shredding and melting. Several factors should be considered along with the security categorisation of the system confidentiality when making sanitisation decisions. The cost versus benefit of a media sanitisation process should be understood prior to a final decision.

iii. Reliability and Viability Issues

Reliability is the level of accuracy an application provides for its intended services, usually dictated by user documentation or application specifications. Reliability is about providing correct results and handling error detection and recovery in order to

avoid failures. More formally, the mean time between failures (MTBF), that is, the average length of time the application runs until a failure occurs, defines reliability. Reliable applications are increasingly critical to customers. Because failure of an application can result in lost data (as well as the possibility of lost business, data for analytics, etc.) and considerable recovery costs, companies are requesting 24x7 reliability in terms of a service level agreement (SLA), commonly known as four 9's or 99.99% availability. Reliability of an application as a whole depends on the reliability of the individual components. Therefore, due to the fact that some components in a system may be related, a failure in one component can affect the reliability of others [14].

Some cloud users always have a concern for the financial stability of the cloud service providers and how long they stay in the market. Though the cloud service providers will remain as they are for a long time, there are chances that they may shut down or they may be acquired by other large service providers. In that case what will happen to the customer's data is always a worry for a cloud user. The data becomes inaccessible or lost. The cloud provider in such a case can divert his service provider but at an extra cost. In such a case how far the data are confidential and secured with the new service provider is one of the great viability issues. Coghead is one example of a cloud platform whose shutdown left the customers in a critical state where they had to rewrite their application to run on a different platform [11].

iv. Geographical and Jurisdictional Issues

The cloud data that are secured in one country are not secured in another country and the user is not sure about where his data are and most of the times it is not disclosed to the user. So the user has to enquire the provider as to the details of data protection laws in the relevant jurisdictions. Currently in the process of trying to harmonise the data laws of its member states, the EU favours very strict protection of privacy, while in America laws such as the US Patriot Act invest government and other agencies with virtually limitless powers to access information including that belonging to companies [4]. Many policy questions will continue to be issues even after the data centre is constructed. The largest challenges to existing providers will likely be tied to issues of security and privacy of the users. Since most data centres are located in the United States, many of these concerns are focused on the USA PATRIOT Act, the Homeland Security Act and other intelligence-gathering instruments like National Security Letters that can be employed by the federal government to compel the release of information. Jurisdiction shopping and the provision of incentives to locate in certain jurisdictions raise several major concerns for users of cloud computing. For example, if certain jurisdictions are too eager for the economic benefits of data centres, they may give away too many legal protections of users and content, granting a great deal of control to the providers. Many organisations are looking to regulation-less safe havens to build their data centres which are always a threat to the consumers. Cloud computing only works if the cloud is massive and contiguous. If geographic and political borders fracture the cloud into smaller groupings, the real advantage of the cloud dissipates into the ether.

v. Regularity Compliance

The security of the data ultimately becomes the responsibility of the cloud user. Most of the times the locations of the data are not disclosed to the user even if the user requires it. Traditional service providers are subjected to external audits and security certifications. Cloud computing providers who refuse to undergo this scrutiny are signalling that customers can only use them for the most trivial functions, according to Gartner [5]. Cloud Computing is not currently subject to specific regulation. However, customers and suppliers of Cloud Computing may be potentially subject to a range of laws: for example, data protection legislation and any relevant industry sector regulations (e.g. financial services and healthcare) .Despite the lack of specific regulation, in certain jurisdictions the provision of Cloud Computing services will require the supplier to obtain a license. For example, in China the provision of SaaS, PaaS or IaaS services will require the supplier to obtain a Type 1 Value Added Telecom Business License. Customers operating in regulated industries such as financial services or healthcare may be subject to even more stringent data protection obligations given the financial value or sensitivity of data such as bank details and medical records. One significant reason for the lack of focus on policy issues about cloud computing is the lack of a political infrastructure that reacts deftly to rapid technological change.

vi. Phishing Issues

Phishing is other major issue the cloud service provider faces. In November 2007, a successful phishing attack compromised contact information on a number of salesforce.com customers, which was then used to send highly targeted phishing emails to salesforce.com users. The phishing breach was cited as an example of why the CRM industry needs greater security for users against such threats as spam. The service has suffered some downtime; during an outage in January 2009 services were unavailable for at least 40 minutes, affecting thousands of businesses. Salesforce.com worked with law enforcement to resolve the problem, but in the meantime it recommended that customers implement a number of security measures in order to cut down on the phisher's chance of succeeding [6].

vii. Transparency Issues

One of the problems regarding the data when they are with a cloud service provider is transparency, and this is not only with cloud computing but anything. It should not be enough for service providers to merely claim that their services are private and secure [18]. Cloud computing vendors could create a self-regulatory code, or they could face regulation from the Government. Recently EPIC has sued Google because they are not transparent on their security policy. Cloud providers should maintain a comprehensive security program and should disclose whether their security efforts meet with security standards.

4 Cloud and Business Growth and the Future

Though cloud computing is a new word and phenomenon for some of the businesses and consumers, business giants are already looking forward to the next big step of

cloud computing. A survey of more than 2,000 companies finds that cloud computing and business analytics solutions that drive growth and innovation are top IT priorities for the midmarket [8].

Companies are now concerned about the growth of business by investing on infrastructure and customer relationship management and no more on cost cutting and cost containment. The survey “Inside the Midmarket: A 2011 Perspective” found that 70 percent of companies are actively pursuing analytics solutions hoping to gain improved customer insight, increased efficiency and better decision-making capabilities. Cloud computing is also growing in the midmarket, with two-thirds of respondents saying they are planning to implement or are currently implementing cloud-based solutions.

5 Future of Cloud

By all ways, cloud computing adoption is growing at a tremendous speed. In Gartner’s 2009 CIO survey, cloud computing is ranked #16 as a business priority. According to Gartner’s latest survey, in one year cloud computing jumped 14 spots up the priority list to #2, behind virtualisation. “Gartner predicts that within two years 80% of Fortune 1000 enterprises will use the cloud at some level.” For many companies, cloud use will be at significant levels -- Gartner has estimated that 30-35% of the IT workload would move to the cloud over the next five years, but the actual rate is outpacing that estimate. Developers and system integrators will reduce as they have not been able to bring any innovations till recently. Low costs and good return on investment will make cloud computing the choice of entrepreneurs. The cloud as a result will grow drastically. This will shift the concept of on-site IT to the cloud environment. Companies will redefine the “C” in “CRM” to mean “Community” rather than “Customer” -- they will build systems that engage their partners and customers in cooperative processes of product and service improvement, rather than building only inward-looking systems for in-house analysis of the world outside the company’s wall [9]. The speed of cloud computing in addressing dramatic increases in demand can be seen in the success of Animoto. This Facebook application allows customers to upload images and music and automatically creates web-based video slideshows from the material. When the application became available on Facebook in March 2008, people liked it and rapidly shared it. By mid-April, the social networking application was spreading like wildfire, and 250,000 people signed up in three days. At one point, as many as 20,000 people tried Animoto every hour. Animoto used the Amazon Web Services infrastructure to manage the scale-up, deploying over 3,000 servers in the three-day bubble to handle the load. And, just as importantly, they were able to release the servers once the spike was past, saving cost for the fledgling company.

The recent IDG Research survey 2011 on the adoption of cloud computing in the enterprise shows there is plenty of consensus around the cloud as shown in Fig.1. [13]

As the IDG Research study shows, cloud adoption is not only growing tremendously but also growing in strategic importance. All indications are that cloud computing will forge ahead in the next three to five years, and that many organisations will try some form of cloud implementation, whether they build their

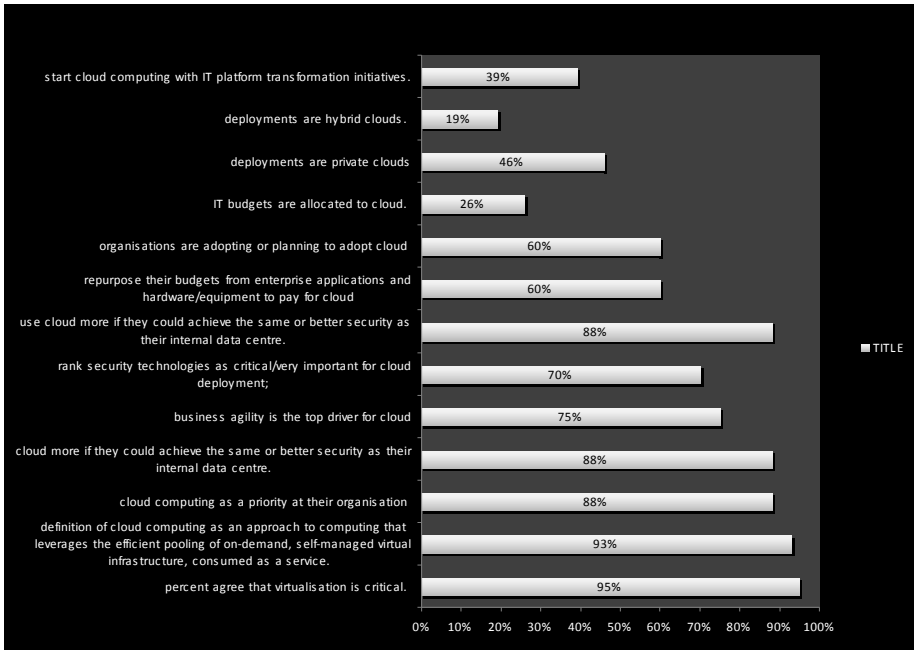


Fig. 1.

own private cloud or move some of their computing requirements out to a public cloud, and with the enterprise hybrid cloud providing great performance and enterprise-level security and quality of service applications, the enterprises will move towards hybrid cloud.

6 Comparison of Cloud System

As cloud’s adoption rate is surging fast, there will be generic as well as specific cloud service providers in large numbers in the hot cloud space. That means that enterprise applications and services will be designed, developed, deployed and delivered from a variety of cloud infrastructures across the globe. Legacy applications will be cloud-enabled and existing business applications will find their new, standard-based, open and cost-effective residence in cloud servers. That is, modernisation and migration will become so casual and common in the near future. Ultimately there will be hundreds of clouds providing thousands of services over the web. As each provider comes out with some attractive services with unique features, consumers want the best services in cheaper costs. Also sophisticated and smart services will be created out of integration across clouds (several services will be linked to create novel services) and thus complex and context-aware services will be created at runtime and will be supplied to users with nominal fee. Even though comparative study has been done by company researchers and academicians on the various cloud service providers available in the market, much has not been done about the study from the point of view of the consumer.

Table 1. Comparison Table of Cloud Service Providers

Service Provider's List	Platform s					Programmin g Languages							Integrat ed Database			
	Linux	Window s	Ma c OS X	CentO S	Othe r	Java	PH P	Per l	C#	Pytho n	Rub y	.Net	MySql	MSSQ L	Orad e	SqlAzur e
Amazon AWS	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No
Windows Azure	No	Yes	No	No	No	No	Yes	No	Yes	No	No	Yes	No	No	No	Yes
RackSpac e	Yes	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No
GoGrid	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	No	No

Table 2. Comparison Table of Cloud Service Providers

Service Provider's List	Service Level Agreement	Free Support	Service Credit for Outage	Incidence Notification	Prepaid Payment	Special Payment Service	Choice of Data Hosting Location	Data Availability After Termination	New User Trail Credentials
Amazon AWS	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Windows Azure	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
RackSpace	Yes	No	No	Yes	Yes	No	No	No	No
GoGrid	Yes	Yes	Yes	No	Yes	No	No	No	Yes

The key considerations dealt in this survey paper are from the perspective of providers and vendors of Cloud Computing Services and specifically for the user. The consumer always finds it difficult to choose a cloud provider which always remains a complicated, calculated and strategic one, which involves the analysis of who-provides-what. So if there is a method which helps the consumer to analyze the various questions that surface in his mind while selecting the cloud it will help the customer in picking the right one. So we try here to do a comparative study of various providers based on IAAS and PAAS as given in Table 1 and 2.

The comparison table has been compiled based on the latest information available on the websites of cloud providers. The contents are subject to change with time, as and when the cloud providers come up with new innovations and features.

7 Conclusion

Cloud computing is the latest computing paradigm that is increasingly popular. Companies such as Microsoft, Google and IBM are trying to promote cloud computing. Difficult application migrations, lack of virtualisation skills, uncertain cost benefits, and risky regulatory environments will lead many IT department managers to dip their toes into the cloud computing pool, but will also keep them from taking the full plunge. Cloud computing models will continue to evolve. Author Nicholas Carr, in his book “The Big Switch,” says: “A hundred years ago, companies stopped generating their own power with steam engines and dynamos and plugged into the newly built electric grid. The cheap power pumped out by electric utilities didn’t just change how businesses operate. It set off a chain reaction of economic and social transformations that brought the modern world into existence. Today, a similar revolution is under way. Hooked up to the Internet’s global computing grid, massive information-processing plants have begun pumping data and software code into our homes and businesses. This time, it is computing that is turning into a utility.

References

1. Singh, T., Vara, P.K.: Smart Metering the Clouds. In: 18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises (2009)
2. <http://www.readwriteweb.com/enterprise/2009/10/forrester-says-we-need-better.php>
3. http://searchcloudcomputing.techtarget.com/sDefinition/0,,sid201_gci1333074,00.html
4. <http://www.computerweekly.com/Articles/2010/01/12/235782/Top-five-cloud-computing-security-issues.htm#1>
5. <http://www.infoworld.com/d/security-central/gartner-seven-cloud-computing-security-risks-853>
6. http://www.pcworld.com/businesscenter/article/139353/salesforcecom_warns_customers_of_phishing_scam.html
7. <http://www.slideshare.net/ConsonusTech/what-is-virtualisation-and-cloud-computing>
8. <http://www.channelinsider.com/c/a/Cloud-Computing/Cloud-Analytics-Will-Drive-Midmarket-Growth-158010/>
9. <http://cloudcomputing.sys-con.com/node/771947>
10. <http://searchcloudcomputing.techtarget.com/definition/hybrid-cloud>
11. Cloudburst as Coghead calls it quits, <http://blogs.zdnet.com/colloboration/?p=349>
12. http://www.psbresearch.com/press_pubs/press_publications_Jan%20_20_2010.html
13. <http://www.idgresearch.com>
14. [http://msdn2.microsoft.com/en-us/library/aa292168\(VS.71\).aspx](http://msdn2.microsoft.com/en-us/library/aa292168(VS.71).aspx)
15. Mikkilineni, R., Sarathy, V.: Cloud Computing and the Lessons from the Past. In: 18th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises (2009)
16. Almulla, S.A., Yeun, C.Y.: Cloud computing security management. In: 2010 Second International Conference on Engineering Systems Management and Its Applications (ICESMA), March 30-April 1, pp. 1–7 (2010)
17. EI-Refaey, M.A., Rizkaa, M.A.: CloudGauge: A Dynamic Cloud and Virtualization Benchmarking Suite. In: 2nd International IEEE Workshop on Colloboration and Cloud Computing (2010)
18. Liver, B., Di Paolo, R., Tice, K.: Privacy in Service Oriented Architectures: SOA Boundary Identity Masking for Enterprises. In: 2010 IEEE 12th Conference on Commerce and Enterprise Computing, CEC, pp. 204–211 (2010)
19. Wang, C., Wang, Q., Ren, K., Lou, W.: Privacy-Preserving Public Auditing for Data Storage Security in Cloud Computing. In: INFOCOM, 2010 Proceedings IEEE, March 14–19, pp. 1–9 (2010)
20. Sotomayor, B., Montero, R.S., Llorente, I.M., Foster, I.: Virtual Infrastructure Management in Private and Hybrid Clouds. IEEE Internet Computing (2009)

Extraction of Urban Road Network by Performing Edge – Aided Classification in Remote Sensing

K. Rajalakshmi^{1,*}, A.L. Dhinaharan², D. Murugan³, and S. Singaravelan⁴

¹ Assistant Professor, Dept. of Computer and Information Technology, Manonmaniam Sundaranar University, India
dhanushkodim@yahoo.com

² Wireless sensor Networks Division, Australia

³ Assistant Professor, Dept. of Computer Science and Engineering, Manonmaniam Sundaranar University, India
dhanushkodim@yahoo.com

⁴ Assistant Professor, Dept. of Computer Science and Engineering, P.S.R.EngineeringCollege, India
singpsr@yahoo.com

Abstract. A new approach for object extraction from high-resolution satellite images is presented in this paper. The new approach integrates image fusion, multi-spectral classification, feature extraction and feature segmentation into the object to improve accuracies. This paper mainly concentrates on road extraction from quick bird MS and Pan images using the proposed approach. Experiments of road extraction with Quick Bird MS and Pan images demonstrate that the proposed approach is effective.

Keywords: Digital, Urban, Object, Edge – Aided, Extraction.

1 Introduction

Road information is essential for cartography, traffic management and planning of urban and industrial areas. For a long time the extraction of road information from aerial and satellite imagery was a labour intensive process. Therefore the automation of road information extraction from the images has been a hot research topic during the past two decades.

Since the recent launch of commercial high resolution satellite (IKONOS and Quick Bird), imagery. The result of road extraction from high-resolution satellite imagery is suitable for large-scale urban applications, such as updating road network of large-scale topographical maps and GIS. Therefore, road extraction from high-resolution satellite imagery holds a great potential.

In this paper, we present a new approach that can be used for urban road network extraction, building extraction and other object extraction.

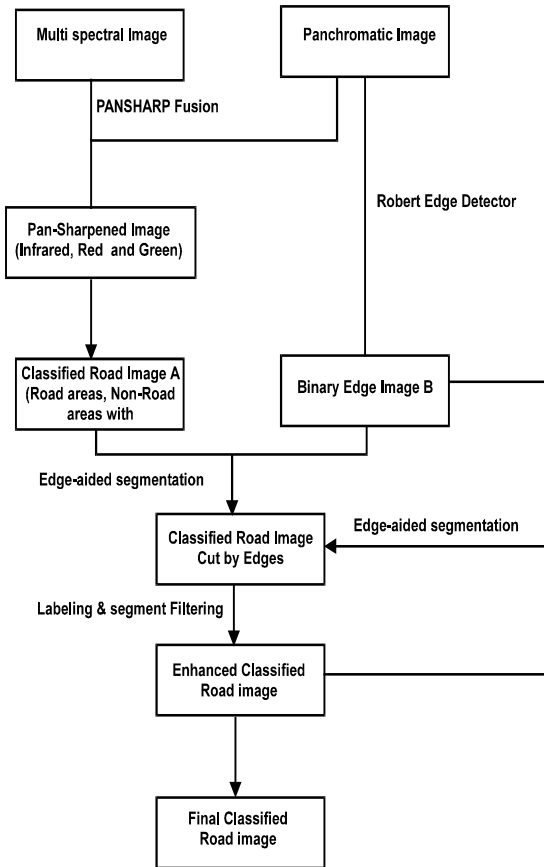
2 The Proposed Approach

The below Figure illustrates the general process of the proposed approach. To overcome the shortcomings at classification of low-resolution MS images, the MS and pan Quick

Bird images are first fused into a pan sharpened MS image. An unsupervised classification is then applied to the Pan sharpened image to obtain a classified road image. And an edge detection approach is applied to the pan image to obtain an edge image.

In the edge aided segmentation the binary edge image from the pan image is employed to segment the classified road image from the Pan-sharpened image. Then a shape – based segmentation and segments filtering algorithm are employed to remove non-road objects.

The whole edge-aided classification process can be iterated to deal with complex road classification results. The individual process of the proposed approach are described in the following sections.



Flow chart of the road extraction procedure

2.1 Pan Sharpening

Pan-sharpening is a technique that produces a high resolution MS image by combining a low resolution MS image with a high – resolution Pan image.

2.2 Classification

The unsupervised clustering method is usually better suited for classifying heterogeneous classes in high resolution satellite images than a supervised classification, the unsupervised fuzzy k mean clustering method was used to classify the Quick Bird images in this study.

2.3 Edge Detection

Sobel, Robert and Canny detectors were compared in this study Robert edge detector can easily achieve a clean and proper edge image from a Quick Bird Pan image.



Fig. 1. Pan-Sharpened MultiquickBird Image

3 Edge Aided Classification:

An edge-aided classification approach was developed to extract accurate road networks from a classified road image.

The edge-aided classification consists of three main processes: Edge-aided segmentation, shape-based segmentation, and segments filtering.

3.1 Edge – Aided Segmentation

In this study, we utilize the edges from the corresponding pan images to separate the non-road objects from the road network. After performing the edge-aided segmentation, those objects connected to road networks are disconnected from the road networks.

3.2 Shape – Based Segmentation

A fast component labeling algorithm is applied to the road images after disconnecting noise, e.g. drive ways and house roofs from the classified road network.

Individual objects, including road networks and noise, are labeled first. They are then segmented according to their size (number of pixels) and shape information (compactness), Resulting in final road networks to be extracted.



Fig. 2. Inverse of Binary Edge Image



Fig. 3. Road Network Cut by Edges

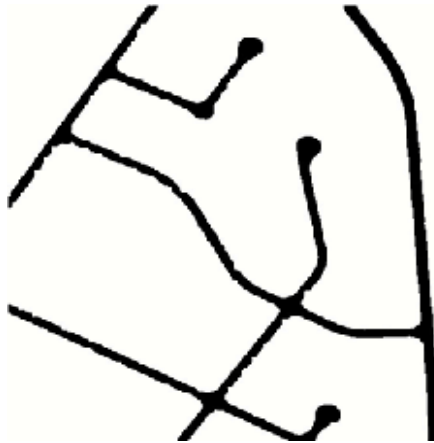


Fig. 4. Extracted Road Network by Edge-Aided Classification

3.3 Segments Filtering

A directional texture detector is developed to distinguish different types of objects according to their texture in different directions. The directional texture detector measures the pixel gray value variance along the central lines in each of four directions are smaller than a certain value, it can be concluded that the object with in

this window is homogeneous. Therefore, this object can be considered as a non-road objects and can be removed.

4 Conclusion

In the previous work they used some basic approach used to road is extracted, but the accuracy is very low. In the existing system they used edge-aided classification method for used to road network cut by edges and extracted road network but the width of road is affected by some parameters like the non-linear edges of the road and gaps found in the center and side of the road. In our proposed system we are going to use effective methods and filters to overcome the previous problem and give an effective road extraction.

Reference

1. Zhang, Y.: Detection of Urban Housing Development by Fusing Multisensor Satellites Data and Performing Spatial Feature Post-Classification. Submitted to the International Journal of Remote Sensing 22(17), 3339–3355 (2001)
2. Zhang, Y.: Texture Integrated Classification of Urban Trees in High-Resolution Color Infrared Imagery. Photogrammetric Engineering & Remote Sensing 67(12), 1359–1365 (2001)
3. Zhang, Y.: A New Automatic Approach for Effectively Fusing Landsat 7 images and IKONOS Images. In: IEEE / IGARSS 2002, Toronto, Canada, June 24–28 (2002)
4. Lec, D.S., Shan, J., Bethel, B.: Classification-guided building detection from IKONOS imagery. In: CD-Proceeding of American Society of Photogrammetry and Remote Sensing, Washington DC (2002)
5. Zhang, C., et al.: Road Network Detection by Mathematical Morphology. In: Proceeding of ISPRS. Workshop “3D Geospatial Data Production: Meeting Application Requirement”, Paris, France, April 7–9, pp. 185–200 (1996)
6. Gong, Wang: Road Network Extraction from High Resolution Airborne Digital Camera Data. Geographic Information Sciences 3(1-2) (December 1997)

Analysis of Texture Patterns in Diagnosing Osteoporosis Using Proximal Femur X-Ray Images

V. Sapthagirivasan¹ and M. Anburajan²

¹ Fulltime Research scholar, ² Professor and Head

Department of Biomedical Engineering,

SRM University, Tamilnadu, India

sapthagiri.ece@gmail.com, anbufelix@yahoo.com

Abstract. Low bone mass pathology otherwise called osteoporosis assessed by BMD quantification. DXA is still referred as, gold standard for BMD assessment. Authenticity of BMD, measured by DXA at anatomic sites such as proximal and spine has incorporated WHO initiation in proposing a tool for diagnosing osteoporosis. In India, osteoporosis is widely evident in post-menopausal women and elder members of both the genders. A direct proportionality exists between degree of mortality and morbidity with hip and spine pathology. Foresight in predicting the risk of fracture is an important physician's goal. Though DXA has been considered as the gold standard for BMD measurement (g/cm^2), the main loophole is the fact that it won't take in to consideration, bone geometry and its micro architecture. The main objective of the present study was to check roughness potential and trabecular bone wavier-ness of the proximal femur that has been sensed by digital x-ray images in post-menopausal osteoporosis evaluation, when DXA has been used as a gold reference for BMD measurement. DXA BMD of the right proximal femur of 26 ($n=26$, mean \pm SD, age = 53.27 ± 14.6 years) south Indian women aged above 25 years was measured. Protocol devised by WHO was adopted in all women. Digital radiograph of right proximal femur was acquired, confirming to medical technicalities. Different methods were used to carry out texture analysis. Relation of BMD with BMI was justified by the results obtained in this study. The present study deciphered the fact that 23% and 27% were affected by osteoporosis and osteopenia respectively. Osteoporotic women exhibited the higher degree of mean roughness, RMS of roughness, mean wavier value and RMS of wavier of neck region compared to normal women as 17%, 15%, 16% and 16% respectively.

Keywords: Osteoporosis, Femur bone, Hip radiograph, Texture analysis, DXA, Computer aided diagnosis (CAD).

1 Introduction

Middle and old aged women population is the most vulnerable factor, being affected by osteoporosis, a multi parametrical pathology. It is often called the "silent disease" because bone loss occurs without any symptoms. Therefore, an osteoporosis diagnosis often occurs after a sudden strain, bash, or fall causes a hip to fracture or a vertebra to collapse. It is one of the prime disorders, concerning healthcare community in India

and across the globe [1]. Bone density tests are useful for diagnosing osteoporosis if a person has already had a suspicious crack/fracture, or for detecting low bone density so that preventive steps can be taken. The most broadly recognized bone mineral density (BMD) test is called a dual-energy x-ray absorptiometry (DXA) test. It is non invasive and painless like having an x-ray, but the radiation exposure level is comparatively less. The BMD can be measured at hip and spine regions. The results of the DXA test are scored in comparison to the bone mineral density (BMD) of young healthy individuals, resulting in a measurement called a T-score. If the T-score is lower than -2.5, they are considered to have osteoporosis and, they are at high risk for a fracture. T-scores between -1.0 and -2.5 are generally considered to show osteopenia (a reduction in bone mass that is not as severe as with osteoporosis). The risk of fractures is generally lower in people with osteopenia when compared to those with osteoporosis. However, if bone loss continues in a person with osteopenia, the risk for fracture increases. There is an additional entry of 2.3 million (0.2%) people to the existing 6 million (5.5%) osteoporotic population. Indian bone statistics revealed the aspect that one in three women; one in eight men is being affected by osteoporotic bone fracture [1-3]. Women exhibit dominant incidents of hip fractures (75%) [4], 25% of the fractures in the hip region target older men (≥ 50 years) [5]. The percentage of hip fracture incidents related to death risk in 50 year old women, during the rest of her life is 2.8% [6]. The concentration, now a days is to recognize patients with frequent fracture risks, rather than to identify the people with osteoporosis assist by BMD alone [7]. There are many aspects of risk fractures, as osteoporosis is just one factor which is expressed in terms of BMD and bone tissue micro architectural deterioration [8]. Fracture risk assessment data should substantiate the factors which can be added to the information that has been provided by the BMD. Numerous quantitative modalities are discovered ranging from simple to sophisticated for BMD or bone mass assessment. The modalities express variations with respect to the anatomical sites, where BMD has to be measured as well as the energy source [9]. The majority of the hip fracture is directly proportional to drops as well as osteoporosis, which paved the way to more incidences of death. The micro structure of the bone, fall direction and severity, family history; femoral neck geometry are some of the predictors of hip fracture risks.

2 Literature Review

A novel low cost screening device adopting principle of conventional radiography used in conjunction with Singh's index by Soontrapa et al [10]. But Singh's index due to its contradictory opinion regarding not well defined grading and cut-off level for osteoporosis, concluded to be unreliable. Nelson et al reported the independent characteristic of vicinity of high fracture risk with hip fracture, in addition to BMD measurement. Therapists as well as physician professionals need the hip fracture as well as BMD data for proper diagnosis of fracture risk [11]. DXR was discovered by Rosholm and fellow researchers. The operating principle of DXR is as follows: Computer measures the cortical thickness of 3 middle metacarpal bones in the hand in a digital x-ray image and it's converted fore arm BMD through a geometrical operation [12]. Prevrhal and fellow researchers, proposed study to evolve two new regions of interest for DXA at hip, one at cortical other at trabecular bone [13].

Sangeetha and co investigators, developed qualitative assessment (wavelength based) of femur bone strength, using digital x-rays. Haar wavelet at 4th level decomposition was used to analyze normal and pathetic femur, and high correlation was demonstrated for abnormal samples [14]. Our previous research work detailed the enhancement of wavelet based qualitative femur bone energy assessment, incorporating digital x-ray images and comparing the same with DXA. There was significance correlation between the results obtained by this method and DXA at 4th level decomposed wavelet coefficients [3]. There was an investigational innovation, originated by Mahadevan.V and co investigators, which signifies trabecular density technique, based on wavelet. The technique also represented anthropometric relationship between height and weight. Changes in shape and size of proximal femur radiograph can be utilized as noticeable references in bone structure estimation [2]. In one of our previous research assets, we suggested morphometry in proximal femur, by the means of which we could spot to catch, bone structural changes [15]. The texture patterns which are observed in the image are the radiographic effect of the underlying anatomical arrangement. The investigation of such natural tree-like structures in digital radiographic images presents special challenges; the surrounding tissue may ambiguous branching patterns. Properties such as topology, spatial distribution of branching, and tortuosity, have been analyzed in the literature and associated with altered function and/or pathology. For example, regional changes in tortuosity of trabecular bone have been used to identify premature stage of osteoporosis development in the human bone [16]. Similarly, studies have shown that the morphology of the ductal network can provide valuable insight to the development of osteoarthritis [17].

The main objective of the present study was to assess trabecular textures potentiality, computed from digital hip radiograph in the reckoning of osteoporosis in Indian women, when it was compared with gold standard technique (DXA-BMD estimation of proximal femur).

3 Materials and Methods

3.1 Study Population

A free osteoporosis testing camp was arranged from the month of august to September 2010 at SRM medical college and hospital. Participants with hepatic, renal, chronic liver, dermatological disorders, hypo- and hyper-thyroidism, malignancy, alcoholism, or receiving medication likely to adversely affect bone health were expelled from the study. 26 women (n=26, mean \pm SD age = 53.27 \pm 14.6 years), who were able to walk about and whole age limit were 25-85 years. None of the participants had the previous history of osteoporotic fracture.

3.2 BMD Evaluation

The right proximal femur BMD was measured by using a DXA, the total body bone densitometer (DPX Prodigy Scanner, GE-Lunar, USA). It sensed BMD at various places of proximal femur sites, such as Neck, Ward's, Trochanter (Greater/Lower), Shaft cortex and Total proximal femur. The study adopted WHO's diagnostic policy

for osteoporosis. Three groups were formed based on DXA measurement of femur neck BMD values obtained by DXA, total women were divided into the following sub-groups: Normal Indian women (n=13, mean ± SD age = 46.2 ±11.03 years and total hip BMD mean ± SD T-Score = -0.5 ± 0.83); Indian women with osteopenia (n= 7, age = 52.1 ±11.82 years and Total hip BMD T-Score = -1.8 ± 0.28); and Indian women with osteoporosis (n=6, age = 70.1 ± 11.8 years and Total hip BMD T-Score = -3.0 ± 0.72).

3.3 Radiographic Evaluation

Digital x-ray machine (Multiphos, Siemens, Germany) was used in procurement of right hip digital radiograph (proximal femur – Anterior Posterior view) in all study Indian women. Total images were acquired with 15 degree internal rotation of femur site. Images attained were in the sort of DICOM. The earned database of image were transformed into bitmap format and normalized with pixel depth of 12 bits per pixel.

The most sensitive region pertaining to bone mass is the ward’s triangle, which is the site that is placed in mid way of the neck region. Figure 1 shows the right femur digital radiograph of 24 years old young normal Indian woman with various ROI used in this paper.

3.4 Texture Analyses

Texture is one of the most important means to classify medical images. It appears to differentiate between normal and abnormal pathologies. Texture may carry substantial information about the structure of physical object. Human’s usually assess texture analysis is required in systems for medical diagnosis. To perform such quantification, mathematically defined texture properties have to be generated by means of texture analysis computer program. Texture analysis of medical images has attracted many investigators particularly for the purpose of developing computer-aided diagnosis (CAD) software systems, which are becoming more and more prevalent due to their ability to increase the precision and accuracy of characterization by radiologists.

The feature vector for texture analysis of a particular image is computed by calculating following features:

a. Average Roughness (S_a):

$$S_a = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} |z(x_k, y_l)| \tag{1}$$

b. The Root Mean Square (RMS) parameter (S_q):

$$S_q = \sqrt{\frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} [z(x_k, y_l)]^2} \tag{2}$$

c. The Surface Skewness (S_{sk}):

$$S_{sk} = \frac{1}{MN S_q^3} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} [(z(x_k, y_l))]^3 \tag{3}$$

d. The Surface Kurtosis (S_{ku}):

$$S_{ku} = \frac{1}{MNS_q^4} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} [z(x_k, y_l)]^4 \quad (4)$$

Where z is image block and M, N are height and width of the image respectively, S_a is average roughness, S_q is root mean square parameter, S_{sk} is surface skewness and S_{ku} is surface kurtosis.

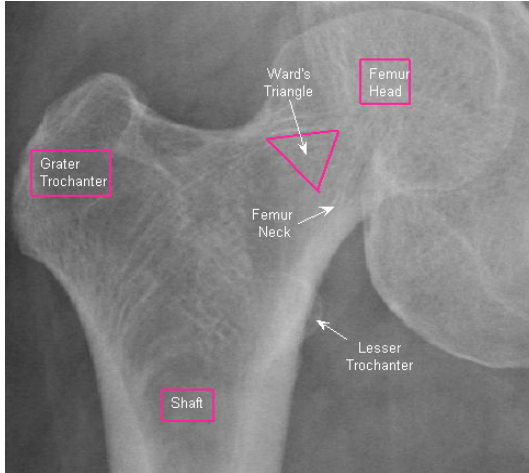


Fig. 1. Digital radiograph with various ROI of right femur

The amplitude properties are described by above parameters (S_a, S_q, S_{sk}, S_{ku}), which give information about the statistical average properties, the shape of the height distribution histogram values and about extreme properties. S_{sk} describes the asymmetry of the height distribution histogram values; S_{ku} describes the 'peakedness' of the surface topography. If $S_{sk} = 0$, a symmetric height distributions (histogram) is indicated, for example a Gaussian like distribution histogram. If $S_{sk} < 0$, it can be a bearing surface with holes and if $S_{sk} > 0$ it can be a flat surface with peaks. If the numerical value of S_{sk} greater than 1.0 may indicate extreme holes or peaks on the surface.

4 Results and Discussion

A database of 26 images (as discussed in section 3.1) was used to evaluate the performance of the proposed approach. Initially all database images were pre-processed by cropping neck region of interest (ROI) and resized them to 256 x 256. Then the method proposed in section 3.4 has applied to the pre-processed image database using MATLAB 7.5 software programming.

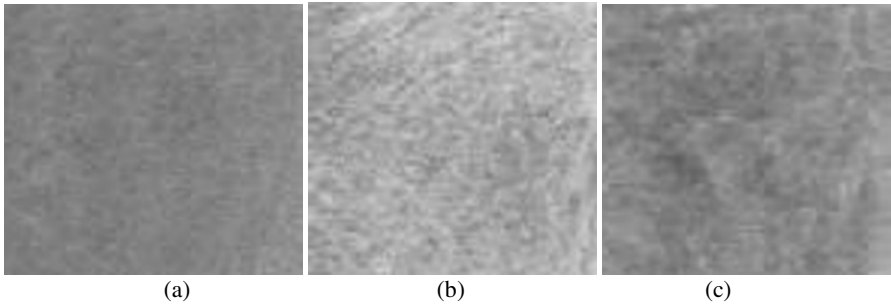


Fig. 2. Neck region of interest in right femur radiograph of (a) Normal 25 years old young Indian woman with BMD of 0.94 g cm^{-2} , (b) 40 years old osteopenic Indian woman with 0.73 g cm^{-2} and (c) 81 years old osteoporotic Indian woman with 0.45 g cm^{-2}

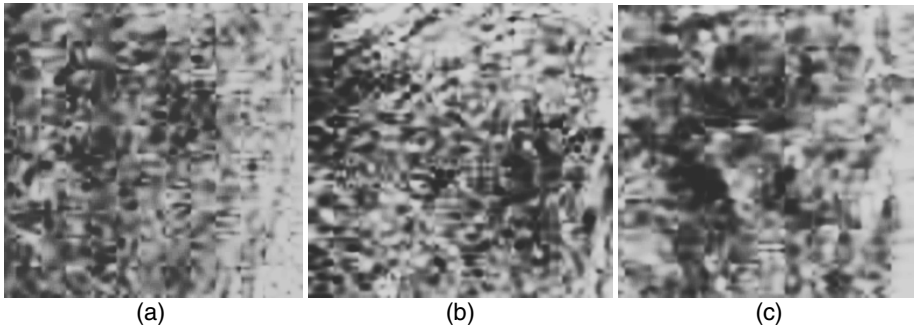


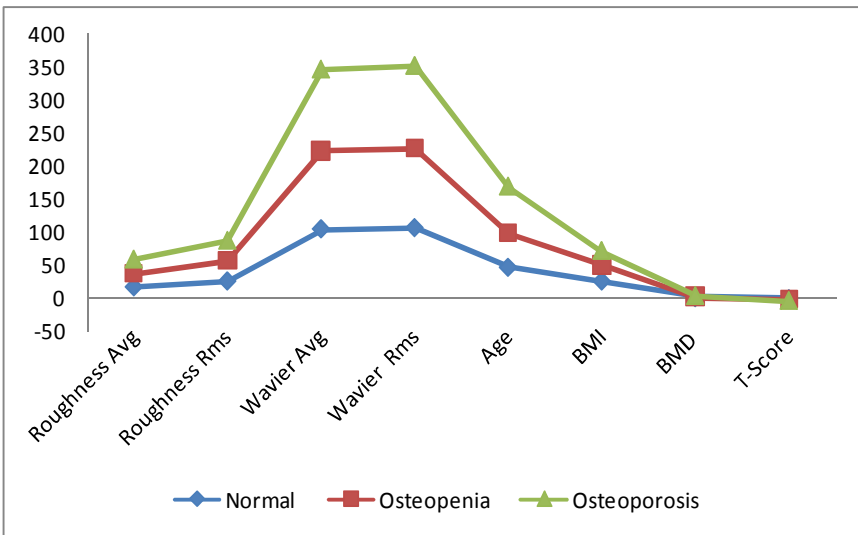
Fig. 3. Enhanced images of neck region of interest in right femur radiograph shown in *Fig.2*

Figure 2(a) shows the pre-processed ROI of 25 years old healthy young Indian woman with BMD of 0.94 g cm^{-2} , Figure 2(b) shows ROI of 40 years old osteopenic Indian woman with 0.73 g cm^{-2} and Figure 2(c) shows ROI of 81 years old osteoporotic Indian woman with 0.45 g cm^{-2} . Table 1 shows the comparison of BMD values obtained by DXA and anthropometric factors such as age, BMI (based on height and weight) for normal, osteopenic and osteoporotic groups. Figure 3(a), (b), (c) displays the enhanced (image intensity equalization) version of images shown in figure 2(a), (b), (c) respectively. Figure 4 shows the comparison graphs of extracted features (average roughness, RMS of roughness, average wavierness and RMS of wavierness) by proposed approach, anthropometric factors (age, BMI-based on height and weight) and BMD obtained by DXA as standard.

BMI, BMD and T-Score shown in Table 1 are gradually decreased for normal, osteopenic and osteoporotic groups, whereas the age is increased. The measured mean femoral neck BMD values in normal Indian women was 0.96 g cm^{-2} , osteopenic Indian women was 0.79 g cm^{-2} , and it's percentage decrease was found to be -22% $[(0.79-0.62)/0.79 \times 100]$, whereas, in osteoporotic Indian women, it was 0.62 g cm^{-2} and it's percentage decrease was found to be -35% $[(0.96-0.62)/0.96 \times 100]$, when comparing to normal Indian women.

Table 1. Comparison of *BMD* obtained by *DXA* and anthropometric factors for normal, osteopenic and osteoporotic groups

Features	Normal N=13	Osteopenia N=7	Osteoporosis N=6	Relationship
Age (years)	46.23 ± 11.1	52 ± 11.8	70 ± 11.8	↑
BMI (Kg/m ²)	25.86 ± 4.7	23.48 ± 2.8	21.3 ± 2.7	↓
BMD (g/cm ²)	0.96 ± 0.1	0.79 ± 0.1	0.62 ± 0.1	↓
T-Score	-0.47 ± 0.8	-1.8 ± 0.3	-2.98 ± 0.72	↓

**Fig. 4.** Comparison graphs of extracted features (average roughness, Rms of roughness, average wavierness and Rms of wavierness) by proposed approach, anthropometric factors (age, *BMI*-based on height and weight) and *BMD* obtained by *DXA* as standard.

The texture results obtained from the experimental studies were shown in Table 2. That is, the texture features extracted were compared with individual groups (normal, osteopenia and osteoporosis) classified by *DXA BMD*. Plotting between features (*x*-axis) and its values (*y*-axis) were shown in figure 4, which shows that clearly mean and RMS of roughness and wavierness values were significantly increasing when *BMD* and *T-Score*'s were decreasing. In Table 2, skewness and kurtosis of roughness values were decreased when *BMD* and *T-Score*'s were decreased, but the significant differences in values were comparatively less.

Figure 5 shows that comparison plot of normalized features shown in figure 4, texture feature and anthropometric feature values were normalized to 0-1 scale and plotted them.

Table 2. Comparison of texture features obtained by proposed method and its relationship with normal, osteopenic and osteoporotic groups (by DXA)

Measured parameters from neck region of digital hip radiograph	Normal N=13	Osteopenia N=7	Osteoporosis N=6	Relation-ship
Roughness Average	17.02 ± 2.9	19.9 ± 3.4	20.54 ± 2.6	↑
Roughness Rms	25.92 ± 4.2	29.55 ± 5.3	30.52 ± 4.2	↑
Roughness Skew	2.06 ± 0.1	1.99 ± 0.1	1.99 ± 0.04	↓
Roughness Kurtosis	5.03 ± 0.5	4.73 ± 0.3	4.72 ± 0.2	↓
Wavier Average	104.1 ± 17.1	118.0 ± 24.9	123.5 ± 20.0	↑
Waiver Rms	105.9 ± 17.2	119.8 ± 25.2	125.38 ± 20.4	↑
Wavier Skew	1.05 ± 0.01	1.04 ± 0.01	1.04 ± 0.01	↓
Waiver Kurtosis	1.12 ± 0.1	1.1 ± 0.01	1.1 ± 0.01	↓

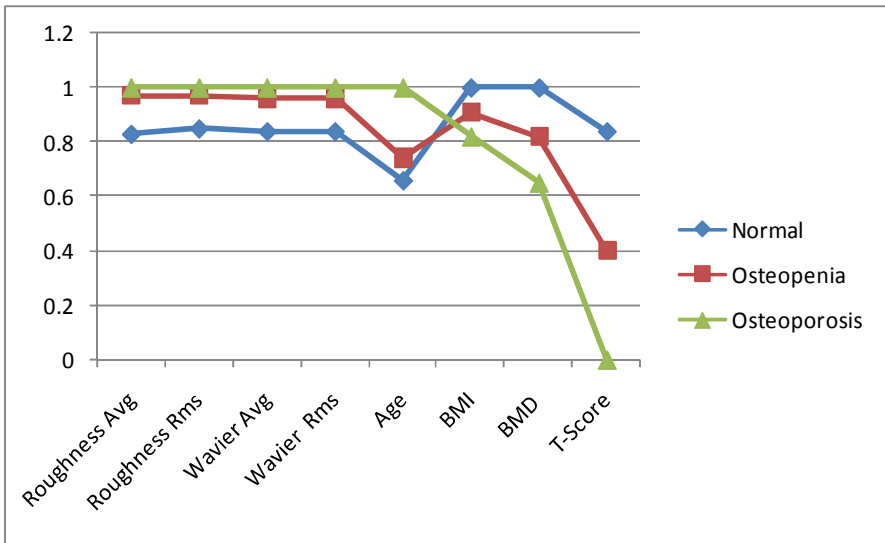


Fig. 5. Comparison of normalized (rescaled to 0-1) features shown in Fig.4

Table 3 categorizes, the values of Pearson’s correlation and its significance of roughness and wavierness (texture features) were compared with anthropometric features (body Weight and calculated BMI). Roughness features shows moderate significance against body weight ($p < 0.05$) and BMI ($p < 0.02$), where as the wavierness parameters shows higher significances ($p < 0.009$, $p < 0.008$) against body weight and BMI respectively.

Table 3. Pearson's correlation values and its significance of texture features Vs anthropometric features

Sno	Parameters	Correlation and Significance	Weight	BMI
1	Roughness Skew	Corr	0.396*	0.461*
		p	0.045	0.018
2	Roughness Kurtosis	Corr	0.420*	0.448*
		p	0.033	0.022
3	Wavier Skew	Corr	0.503**	0.488*
		p	0.009	0.011
4	Waiver Kurtosis	Corr	0.518**	0.510**
		p	0.007	0.008

* Moderately significant,

** Highly significant

5 Conclusion

This study reveals the fact that 23% of the recruited female subjects under study were found to have osteoporotic and 27% were found to be osteopenic based on BMD obtained by DXA. Texture analysis has successfully applied to the trabecular pattern recorded in the radiograph of right proximal femur. Osteoporotic women exhibited the higher degree of mean roughness, RMS of roughness, mean wavier value and RMS of wavier of neck region compared to normal women as 17%, 15%, 16% and 16% respectively. Roughness features shows moderate significance ($p < 0.05$), whereas the wavierness shows high significance ($p < 0.009$) against body weight and calculated BMI in overall groups and which shows higher similarity with the work done by Huber M.B. et al [18] and Corroller T.L [19]. The extracted features from trabecular pattern were able to give information about the quality of the bones for the assessment of osteoporosis. The fractal dimension using box counting algorithm also has a significant correlation to bone quality assessment using trabecular energy with BMD and these computerized algorithms, it is possible to use them for screening purpose as radiographic facilities are available all over the country. The early detection of osteoporosis by the change in trabecular bone assessed using these algorithms will become a significant contribution to improve the quality of healthcare.

Acknowledgments. Authors wish to thank Department of Orthopedic and Traumatology, Faculty of Medicine, SRM University, for suggesting radiographic guidance and also thankful to Arthi Scans Chennai, India for their support for providing continuous support.

References

1. International Osteoporosis Foundation, <http://www.iofbonehealth.org/facts-and-statistics.html> (accessed on April 23, 2011)

2. Mahadevan, V., Sapthagirivasan, V.: Information Processing of Medical Images for the Detection of Osteoporosis in Hip Region of Interest. *Intl. J. Innovation, Management and Technology* 1(5), 516–520 (2010)
3. Sapthagirivasan, V., Anburajan, M., Mahadevan, V.: Bone trabecular microstructural properties evaluation of human femur using mechanical testing, digital X-ray and DXA. In: *IEEE Intl. Conf. on Software and Computing Technology*, Kunming, China. vol. 1, pp. 1–6 (2010)
4. Jordan, K.M., Cooper, C.: Epidemiology of osteoporosis. *Best Pract. Res. Clin. Rheumatology* 16(5), 795–806 (2002)
5. Cooper, C., Campion, G., Melton, L.J.: Hip fractures in the elderly a world-wide projection. *Osteoporosis Int.* 2, 285–289 (1992)
6. Cummings, R.S., Black, D.M., Rubin, S.M.: Lifetime risk of hip, Colle's, or vertebral fracture and coronary heart disease among white postmenopausal women. *Arch. Intern. Med.* 149(11), 2445–2448 (1989)
7. Siris, E., Delmas, P.D.: Assessment of 10-year absolute fracture risk: a new paradigm with worldwide application. *Osteoporosis Int.* 19, 383–384 (2008)
8. Kanis, J.A., Johnell, O., Oden, A., Johansson, H.: FRAX and the assessment of fracture probability in men and women from the UK. *Osteoporosis Int.* 19, 385–397 (2008)
9. Tothil, P.: Methods of bone mineral measurement-review article. *Phys. Med. Biol.* 34, 543–572 (1989)
10. Soontrapa, S., Srinakaran, J., Chowchuene, P.: Singh index screening for femoral neck osteoporosis. *Khon Kaen University* (2005)
11. Nelson, D.A., Beck, T.J., Wu, G., Lewis, C.E., Bassford, T., Cauley, J.A., LeBoff, M.S., Going, S.B., Chen, Z.: Ethnic differences in femur geometry in the women's health initiative observational study. *Osteoporosis Int.* (July 2010)
12. Rosholm, A., Hyldstrup, L., Backsgaard, L., Grunkin, M., Thodberg, H.H.: Estimation of bone mineral density by digital X-ray radiogrammetry: theoretical background and clinical testing. *Osteoporosis Int.* 12(11), 961–970 (2001)
13. Prevrhal, S., Meta, M., Genant, H.K.: Two new regions of interest to evaluate separately cortical and trabecular BMD in the proximal femur using DXA. *Osteoporosis Int.* 15, 12–19 (2004)
14. Sangeetha, S., Christopher, J.J., Ramakrishnan, S.: Wavelet based qualitative assessment of femur bone strength using radiographic imaging. *Intl. J. Biological and Life Science* 3(4), 276–279 (2007)
15. Shankar, N., Sapthagirivasan, V., Vijay, A., Kirthika, K., Anburajan, M.: Evaluation of Osteoporosis Using Radiographic Hip Geometry, Compared with Dual energy x-ray absorptiometry (DXA) as the Standard. In: *IEEE Intl. Conf. on Systems in Medicine and Biology*, pp. 272–277 (2010)
16. Szczypinski, P.M., Strzelecki, M., Materka, A., Klepaczko, A.: Mazda – A Software Package For Image Texture Analysis. *Computer Methods and Programs in Biomedicine* 94(1), 66–76 (2009)
17. Woloszynski, T., Podsiadlo, P., Stachowiak, G.: A Signature Dissimilarity Measure for Trabecular Bone Texture in Knee Radiographs. *Medical Physics* 37(5) (2010)
18. Huber, M., Carballido-Gamio, B., Majumdar, J., Link, S.: Development and testing of texture discriminators for the analysis of trabecular bone in proximal femur radiographs. *Med. Phys.* 36(11), 5089–5099 (2009)
19. Le Corroller, T., Halgrin, J., Pithioux, M., Guenoun, D., Chabrand, P., Champsaur, P.: Combination of texture analysis and bone mineral density improves the prediction of fracture load in human femurs. *Osteoporosis Int.* doi:10.1007/s00198-011-1703-1

Energy Efficient Data Aggregation in Cognitive Sensor Networks

Sumalatha Ramachandran¹, Aswin Kumar Gopi², Giridara Varma Elumalai²,
and Murugesan Chellapa²

¹Assistant Professor, ²Research Student,
Department Of Information Technology, Anna University

Abstract. In this paper, a protocol for energy efficient data aggregation process is proposed. A virtual infrastructure is considered where the entire geographical area of interest is divided into grid based clusters having a representative member in each of the cluster. The output from the sensors is made to pass through a context aware system to ensure the validity of the sensor data. Redundancy in the above valid data can be of two types – Intra cluster and Inter Cluster. Based on the type of redundancy, redundancy elimination is performed at header nodes or at intermediate nodes thus facilitating energy efficient aggregation of data from source node to sink.

1 Introduction

In wireless sensor networks, sensors are employed in the area of interest to gather data or to monitor the environment. Sensors can be deployed in a random fashion or can be placed in strategic locations. The data collected by each node is forwarded to the sink (destination) through a process known as “**Data Dissemination**”. However the sensors in wireless sensor networks are battery powered due to which it has limited lifetime. So the data dissemination policy has to be framed in such a way that the energy consumption is optimal.

2 Related Work

The entire geographical area is divided into grid based clusters[1]. Each cluster has a header node which acts as a representative node of that cluster (Fig 1). The header node is responsible for the location update of the mobile sink groups. The header node in each cluster is elected using backward timer. The access node (sink) queries the source node for data. The data dissemination process from source node to access node (sink) is exploited using header to header forwarding[2].

Whenever an event occurs in the monitoring environment, the sensor node detects it and forwards the relevant data to the sink node for processing and to take the necessary actions[3][4]. Eq.,1 below shows the relationship between sensor lifetime and battery power as, lifetime is directly proportional to the battery power remaining.

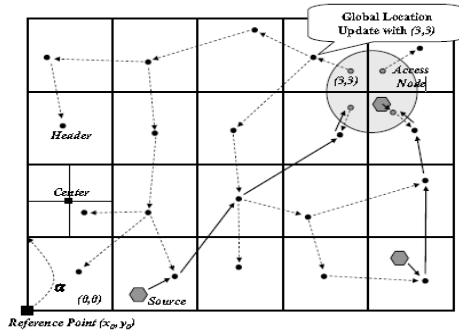


Fig. 1. Cluster based virtual infrastructure

Lifetime of a sensor node	α	Battery power of the sensor node
----------------------------------	----------	---

Eq.,1

Battery power is consumed when data is transmitted. Hence by Eq.2, more the amount of data transmitted, more the power consumed. More the power consumed, less the life time of the sensor nodes.

Amount of data transmitted	α	Amount of power consumed
-----------------------------------	----------	---------------------------------

Eq.,2

So the data **dissemination** process has to be efficient in order to reduce the amount of data transmitted and increase the lifetime of the sensor nodes.

3 Proposed Work

In the above virtual infrastructure when a member sink queries two or more source node for data there is a high probability of data redundancy in data dissemination process[8]. Hence if redundant data are to be transmitted again and again, the amount of data transmitted increases which ultimately results in high consumption of power and bandwidth.

3.1 Types of Redundancy in the Above Infrastructure

Intra-Cluster redundancy: When the source nodes lie in the same cluster.

Inter-Cluster redundancy: When the source nodes lie in adjacent or neighboring clusters.

Due to the above condition the lifetime of the sensor node decreases reducing the reliability of the network.

3.2 Redundancy Elimination

The main objective of the project is to increase the lifetime of the sensor by framing an efficient data dissemination policy which involves identifying redundant data and eliminating the same before the data reaches the sink node.

3.2.1 Eliminating Intra-Cluster Redundancy

When the source nodes lie in the same cluster the data to be transmitted is compressed at the source nodes. The compressed data is then forwarded to the header node where redundancy elimination takes place (Fig 2). Then the data dissemination is brought about by header to header forwarding until the data reaches the access node (Sink).

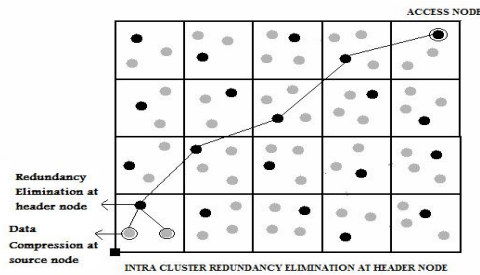


Fig. 2. Intra- Cluster Redundancy Elimination

3.2.2 Eliminating Inter-Cluster Redundancy

When the source nodes lie in adjacent clusters the shortest path from header node to access node for each source node is found. These paths tend to converge at a header node present in some other cluster. Elimination of redundant data is performed at this header node and compression at their respective source header (Fig 3).

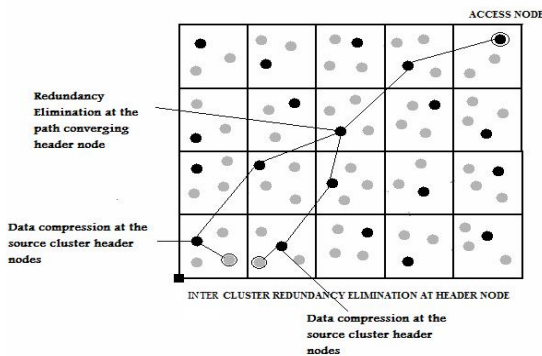


Fig. 3. Inter-Cluster Redundancy Elimination

3.3 Module Description

The proposed architecture of Redundancy Eliminated Data Dissemination consists of the following modules,

1. Cluster Formation
2. Dynamic Topology Management
3. Header Node Election
4. Data Compression
5. Context Aware System
6. Finding Shortest Node Converging Path
7. Redundancy Elimination

3.4 Architecture of REDD

Each node is deployed in the area of interest to observe the environment around and must belong to a cluster defined by its geographic axes. The nodes that fall within a particular cluster form a group under a representative node called Header Node. This is done by **Cluster Formation** and **Header Election** modules. Since the sensor nodes are mobile, the position of each sensor node keeps on changing. So the sensor node associates itself with a particular cluster from time to time with the help of **Dynamic Topology Management** module.

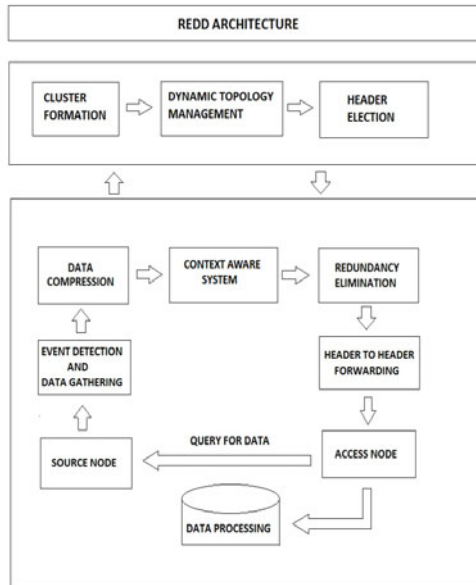


Fig. 4. System Architecture

The job of detecting events nearby is done by the sensors at the physical level. These events are then combined together and stored locally. The sink node queries for data of interest from the source nodes through the **Querying** module. The query is then forwarded to the access nodes by **Header to Header** forwarding [9]. For this forwarding, the Shortest Path and the Shortest Converging Path are found by the sink. Based on the destination of the sinks, the corresponding path is used. While replying, data is compressed at the cluster head. The compressed output from the sensor is made to pass through a **Context Aware System** to ensure reliability of the sensor data. Redundancy in the valid data is eliminated either at the header node or at some intermediate junction node which is mentioned in the query by the sink node. The redundancy eliminated reply from the access nodes is then sent to the sink by Header to Header forwarding in the same way as the query.

On receiving the data, the **Data Processing** module of the sink node processes the data and performs the necessary action. Finally, the performance of the REDD protocol is analyzed graphically and the results are tabulated

4 Implementation Issues

4.1 Algorithm: Cluster Formation

Form_Cluster(Node temp)

```

1.  $x_p \leftarrow \text{getXPosition}()$ ;
2.  $y_p \leftarrow \text{getYPosition}()$ ;
3.  $\text{RFID}_{\text{node}} \leftarrow \text{getRFID}(x_p, y_p)$ ;
4.  $x_c \leftarrow (x_p / G_{\text{size}}) * G_{\text{size}} + G_{\text{size}} / 2$ ;
5.  $y_c \leftarrow (y_p / G_{\text{size}}) * G_{\text{size}} + G_{\text{size}} / 2$ ;
6.  $\text{RFID}_{\text{header}} \leftarrow \text{getRFID}(x_c, y_c)$ ;
7. if ( $\text{RFID}_{\text{header}} == \text{NULL}$ )
8.   setHeader( $\text{RFID}_{\text{node}}$ );
9. else
10.  sendMsg( $\text{RFID}_{\text{header}}$ );
11.  wait();
12.  if(receiveMsg())
13.    setHeader( $\text{RFID}_{\text{header}}$ );
14.    joinMember( $\text{RFID}_{\text{node}}$ );
15.  endif
16. endif
17. end

```

Note:

1. When a node is created it contains the following information
 - a) RFID – Unique identity of the node
 - b) Location information (X,Y co-ordinates)

2. To find the grid to which the node belongs following formula is used

$$x_c \leftarrow (x_p / G_{size}) * G_{size} + G_{size} / 2 ;$$

$$y_c \leftarrow (y_p / G_{size}) * G_{size} + G_{size} / 2 ;$$

where x_c, y_c are the co-ordinates of header and G_{size} is the grid size

The above algorithm gets the position of the node from the GPS device, calculates the grid center and joins the grid header if any. If there is no header yet in the grid, it becomes the header node.

4.2 Algorithm: Location Update (Dynamic Topology Management)

Update_Location(Node node)

```

1.  xold ← getXPosition();
2.  yold ← getYPosition();
3.  RFIDnode ← getRFID(xold, yold);
4.  xoc ← ( xold / Gsize ) * Gsize + Gsize / 2 ;
5.  yoc ← ( yold / Gsize ) * Gsize + Gsize / 2 ;
6.  RFIDcurrentheader ← getRFID(xoc, yoc);
7.  xnew ← setXPosition();
8.  ynew ← setYPosition();
9.  xnc ← ( xnew / Gsize ) * Gsize + Gsize / 2 ;
10. ync ← ( ynew / Gsize ) * Gsize + Gsize / 2 ;
11. RFIDnewheader ← getRFID(xnc, ync);
12.   if (RFIDnewheader == RFIDcurrentheader )
13.       return;
14.   else if (RFIDnewheader != RFIDcurrentheader)
15.       sendMsg(RFIDnewheader)
16.       wait();
17.       if(receiveMsg())
18.           setHeader(RFIDnewheader);
19.           joinMember(RFIDnode);
20.       endif
21.   endif
21. End

```

Note:

1. If a node moves within in its grid there is no change in topology.
2. If a member node of one grid moves to another it registers itself with the header of the new grid.

Whenever a node moves, it calculates its new grid center and checks with the old center. If they are not the same, it joins under the new header node and unregisters from the old header. If it is a header itself it initiates election process in the old grid.

4.3 Algorithm: Header Node Election

Elect_Header()

```

1. Initialize node[i]battery ← 100 for all I;
2. if (i%5==0)
3.   GRID(i/5)/Gsize, (i/5)%Gsize header ← node[i/5*5];
4. end if
5. if (onBatteryChange())
6.   if (Cur_HeaderBattery<100&&Cur_HeaderBattery>50)
7.     continue;
8.   else
9.     Old_Header ← Cur_Header;
10.    Cur_Header ← (Cur_Header.Members[] →next);
11.    Cur_Header.Members[].Join ← Old_Header;
12.  end if
13. end if
14. end

```

If the current header has moved out of the grid or if its battery power has reduced below the threshold, it starts the election process and notifies the first node among its members. That node, if it has enough power, becomes the new header and advertises its election as the new header to other member nodes.

4.4 Algorithm: Sensor Data Forwarding

Forward(Graph, Source)

```

1. for each vertex v in Graph:
2.   dist[v] := infinity ;
3.   previous[v] := undefined ;
4. end for ;
5. dist[source] := 0 ;
6. Q := the set of all nodes in Graph ;
7. while Q is not empty:
8.   u := vertex in Q with smallest dist[] ;
9.   if dist[u] = infinity:
10.    break ;
11.  endif ;
12.  remove u from Q ;
13.  for each neighbor v of u:
14.    alt := dist[u] + dist_between(u, v) ;
15.    if alt < dist[v]:
16.      dist[v] := alt ;
17.      previous[v] := u ;
18.    endif ;
19.  end for ;
20. end while ;
21. return dist[] ;
22. end

```

When a node wants to send data to some other node, it first calculates the shortest path to the destination graphically and then forwards the data to the next header node in the path along with information regarding the path by Header to Header forwarding [9]. The next header node on receiving the data forwards it to the next down lower in the order and so on.

Thus, all the above algorithms are used in the implementation of the REDD architecture.

4.5 Context Aware System

In wireless sensor networks, the reliability is ensured by making use of redundant data. But this is facilitated at the cost of excessive power consumption. Since we are going to remove the redundancy, we have to ensure that the unique data has to be a reliable one [11][13]. To ensure that we use a context aware system as shown in Fig 5. The context aware architecture consists of the following modules

Sensor Output - Contains raw data about temperature and pressure in byte stream.

Context Interpreter – Extracts the context (temperature | Pressure) from the raw sensor data.

Context Retriever – Retrieves the context that has to be compared with a set of rules.

Rule Engine – Provides the set of rules as described in Table 1.

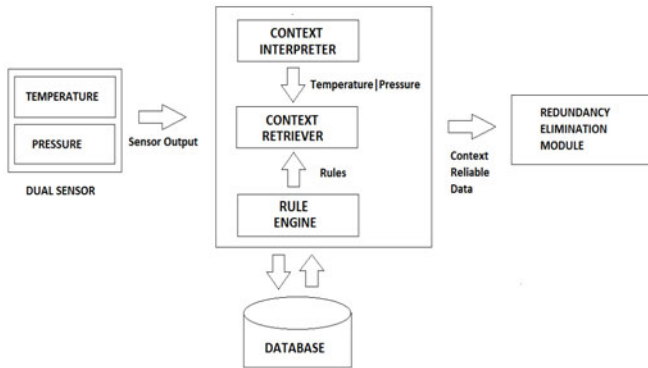


Fig. 5. System Architecture

Java context awareness framework (JCAF) [10] is used to implement the context aware system. The compressed output from the sensor data is given as input to this module. The temperature and pressure context are extracted from the compressed output and then it is checked for validity against the rules defined in the rule engine. If the context is valid the data is forwarded to redundancy elimination module else it is suppressed.

Table 1. Initial Set of Rules

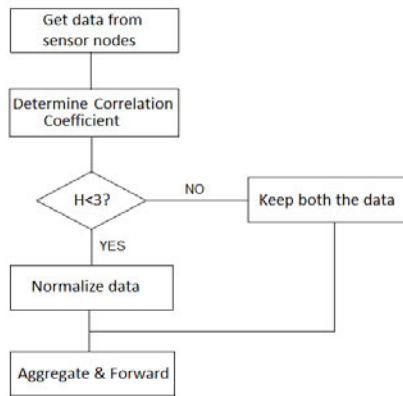
S.No	Condition	Action
1	Temperature < UTmin	Discard
2	Temperature > UTmax	Discard
3	UTmin < Temperature < UTmax	Allow
4	Pressure < UPmin	Discard
5	Pressure > UPmax	Discard
6	UPmin < Pressure < UPmax	Allow

UTmin & UTmax – Minimum and Maximum allowable temperatures
 UPmin & UPmax - Minimum and Maximum allowable Pressure

4.5 Redundancy Elimination

One technique used to decrease the number of redundant messages transmitted and thus pro-long the network lifetime is data aggregation. REDD protocol's redundancy elimination algorithm is based on a single value called the **correlation coefficient** to represent the whole set of readings recorded by all the nodes in the sensor field.

The value of the correlation coefficient (H) ranges from 1 to 10. $H = 1$ is for strong correlation of data and as the correlation coefficient increases, the degree of correlation between data decreases. The following flowchart (Fig 5) describes the above procedure

**Fig. 6.** Redundancy Elimination

5 Performance Analysis

The initial network deployment consists of 3x3 grids with 5 nodes in each grid respectively. The battery level of each sensor is initially set to maximum (100%). The transmission cost is equal to one unit per bit. (Table 2)

The following graph (Fig 7) represents the validity ratio which is defined as the ratio between number of valid transmissions and number of total transmissions. It can be inferred that as the number of transmission increases the validity ratio decreases.

As the total number of transmissions increases, the number of suppressed invalid transmissions also increases. Hence resulting in more power saving.

Table 2. Simulation Scenario

No of transmissions	No of valid transmissions	No of invalid transmissions	No of redundant transmissions	No of unique transmissions
100	95	5	2	93
200	193	7	6	187
300	287	13	14	273
400	375	25	27	348
500	468	32	36	432

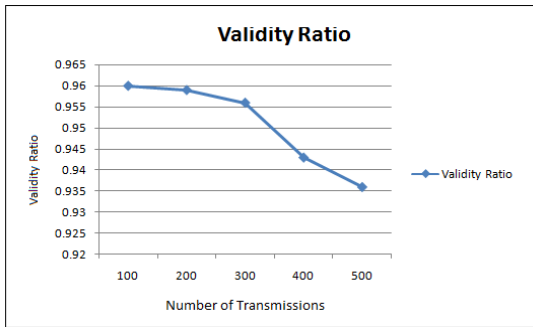


Fig. 7. Validity Ratio

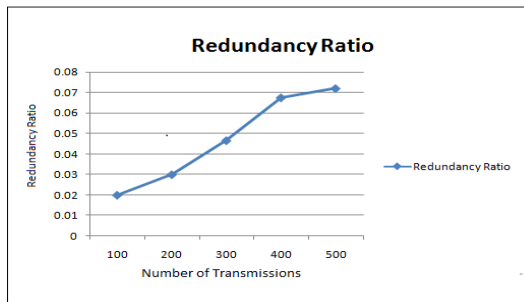


Fig. 8. Redundancy Ratio

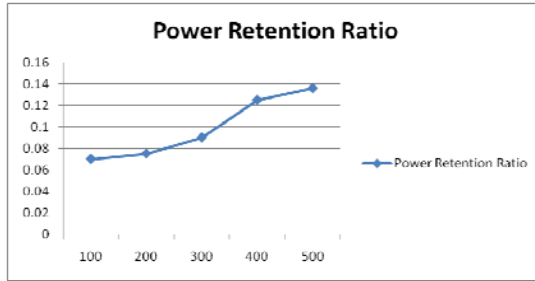


Fig. 9. Power retention Ratio

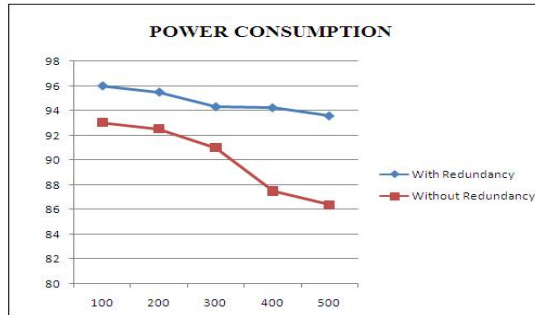


Fig. 10. Power Consumption

The redundancy ratio which is defined as the ratio between the number of redundant transmissions and total number of transmissions shows that as the number of transmissions increase the redundancy also increases. This is depicted in the following graph (Fig 8). In the graph below it can be seen that if the number of transmission packets is 500 the redundancy ratio is almost 0.07 which is relatively high and increases the amount of power wasted due to redundant transmissions. In order to reduce the amount of power wasted there is a necessity for redundancy elimination process.

Power retention ratio is the ratio between number of suppressed transmissions and total number of transmissions. The power retention ratio graph below (Fig 9) depicts that when considering only unique transmissions after the redundancy elimination process, the amount of power saved is relatively high.

The power consumption graph shown in (Fig.10) implies that the amount of power consumed with redundant transmissions is higher when compared to the power consumed after eliminating redundant transmissions, thereby proving the effectiveness of the REDD architecture.

References

1. Park, S., Lee, E., Jin, M.S., Kim, S.H.: Novel Strategy for data dissemination to mobile sinks in wireless sensor networks. *IEEE Communication Letters* 14(3), 202–204 (2010)
2. Viana, A.C., Ziviani, A., Friedman, R.: Decoupling Data Dissemination from Mobile Sink's Trajectory in Wireless Sensor Networks. *IEEE Communication Letters* 13(3) (March 2009)

3. Hamida, E., Chelius, G.: Strategies for data dissemination to mobile sinks in wireless sensor networks. *IEEE Wireless Commun.* 15(6), 31–37 (2008)
4. Machado, Goussevskaia, Mini, Rezende, Loureiro, Mateus, Nogueira: Data Dissemination in Autonomic Wireless Sensor Networks. *IEEE Wireless Commun.* 23(12) (December 2005)
5. Song, L., Hatzinakos, D.: Architecture of Wireless Sensor Networks With Mobile Sinks: Sparsely Deployed Sensors. *IEEE Transactions on Vehicular Tech.* 56(4) (July 2007)
6. Ye, F., Luo, H., Cheng, J., Lu, S., Zhang, L.: A two-tier data dissemination model for large-scale wireless sensor networks. In: *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)* (September 2002)
7. Park, S., Lee, D., Lee, E., Yu, F., Choi, Y., Kim, S.: A communication architecture to reflect user mobility issue in wireless sensor fields. In: *Proc. IEEE Wireless Communications and Networking Conference, WCNC* (March 2007)
8. Hamida, E.B., Chelius, G.: Line-based data dissemination protocol for wireless sensor networks with mobile sink. In: *Proc. IEEE International Conference on Communications, ICC* (May 2008)
9. Yu, Y., Govindan, R., Estrin, D.: Geographical and Energy Aware Routing: A Recursive Data Dissemination Protocol for Wireless Sensor Networks. Technical Report UCLA/CSD-TR-01-0023, UCLA Computer Science Dept. (May 2001)
10. Bardram, J.E.: The Java Context Awareness Framework (JCAF) – A Service Infrastructure and Programming Framework for Context-Aware Applications
11. Baldauf, M., Dustdar, S., Rosenberg, F.: A Survey On Context-Aware Systems. *International Journal of Ad Hoc and Ubiquitous Computing* 2(4), 63–277 (2007)
12. Kjaer, K.E.: A Survey of Context-Aware Middleware. In: *Proceedings of the IASTED Software Engineering Conference* (2007)
13. Singh, A., Conway, M.: Survey of Context aware Frameworks – Analysis and Criticism. UNC-Chapel Hill ITS Version: 1 (2006)

VLSI Architecture for Compressed Domain Video Watermarking

N. Mohankumar, M. Nirmala Devi, D. Badari Nath, and Arun Scaria

VLSI Design Research Group
Department of Electronics and Communication Engineering, School of Engineering
Amrita Vishwa Vidyapeetham, Coimbatore, India
{n_mohankumar, m_nirmala}@cb.amrita.edu,
badri.urs2007@gmail.com, arunscaria86@yahoo.com

Abstract. Digital watermarking has become very important for protecting the authenticity of multimedia objects as they become easier to copy, exchange, and modify due to the large diffusion of powerful personal computers. The video has been utilized in a variety of applications such as video editing, Internet video distribution, wireless video communications etc. Some of these applications are likely to get great benefit from video watermarking technology. Main objective of this research is to design robust perceptual video watermarking targeted at achieving better performance and reliability. Due to its robust nature, Discrete Cosine Transform (DCT) watermarking was chosen in this work to accomplish video copyright protection. The watermark is inserted in the video stream during compression, resulting in an optimized compression/watermarking algorithm and system.

Keywords: Digital video Watermarking, Copyright Protection, video compression, Multimedia Content Protection.

1 Introduction

With the growth of broadband and multimedia content flowing across networks, a major issue is to develop a secure technology that protects the content. As benefits of the internet come with lack of security, research on public key video watermarking to protect multimedia content is required [1].

Watermarking is the process that embeds data called watermark or digital signature into a multimedia object such that watermark can be detected or extracted later to make an assertion about the object. A simple water marking scheme is shown in Fig.1. The main purpose of digital watermarking is to embed information imperceptibly and robustly in the host data. Typically the watermark contains information about the origin, ownership, destination, copy control, transaction etc. Potential applications of digital watermarking include transaction tracking, copy control, authentication, legacy system enhancement and database linking.

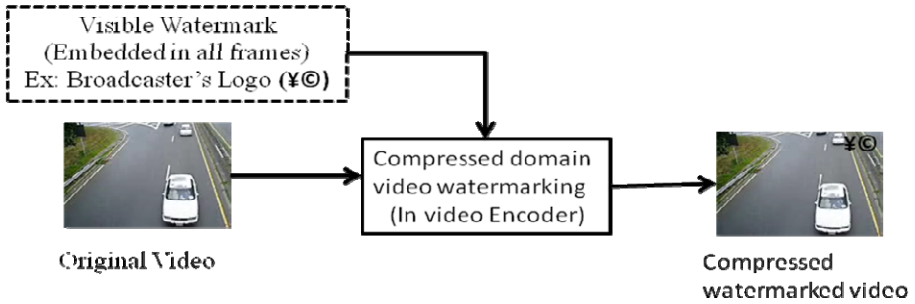


Fig. 1. Visible watermark insertion in compressed domain watermarking

Growing popularity of video based applications such as Internet multimedia, wireless video, personal video recorders, video-on-demand, set-top box, videophone and videoconferencing have a demand for video watermarking. A video watermark is a pattern embedded into compressed domain video content [7]. The output is real-time compressed video with the visible watermark or logo embedded. By embedding a unique watermark into video material at the time of production or broadcasting, content owners can identify copies of their material and whether they are legal or not.

2 Hardware versus Software

Software-based watermarking [2] provides the following:

- The algorithm's operations are performed as code running on a microprocessor.
- Abstraction of the implementation from any hardware details.
- Availability of software tools to aid in realizing various data operations.
- Limited means of improving area and time complexity (speed) of the implementation.

Hardware-based watermarking operations are fully implemented in custom-designed circuitry and consume less area and power. In consumer electronic devices, a hardware watermarking is more economical due to addition of small dedicated area of silicon watermarking component.

3 Research Review

In conventional video watermarking, a compressed video stream is first decompressed into standard video, a watermark signal is then embedded into the video signal, and finally, the watermarked video is recompressed. This technique requires fully decompressing and recompressing the video stream. This method consumes a lot of computer processing time, thereby making it computationally very intensive.

L.De Strycker et.al. [1] proposed a real-time invisible digital watermark embedder and detector for television broadcast monitoring. Broadcast material is pre-encoded

with an invisible, unique watermark identifier A TriMedia DSP processor with 4 BOPS (billion operations per second) is used for a software implementation. This algorithm is not useful for logo insertion. N. J. Mathai et.al. [2] Proposed a hardware implementation aspects of digital watermarking through Just Another Watermarking System (JAWS). Here, time and area constraints were considered and the relation between algorithmic features and implementation cost is analyzed. A. Garimella et.al. [3] proposed a VLSI architecture implementation for Fragile watermarking in spatial domain. The ASIC design uses 0.13 μ CMOS technology. The area of the chip is 3,453x3,453 μm^2 and the power consumption is 37.6 μ W. Y. C. Fan et. al [4] proposed an adaptive DWT (discrete wavelet transforms) based visible watermarking design. Due to high computation penalty in DWT which is not suitable for real-time applications, DCT is chosen in this work to accomplish MPEG-4 video copyright protection for real-time applications.

S. Biswas, S. R. Das, and E. M. Petriu [5] proposed a compressed video watermarking which embeds several binary images, decomposed from a single watermark image, into different scenes of a video sequence. The spatial spread spectrum watermark is embedded directly into the compressed bit streams by modifying discrete cosine transform (DCT) coefficients.

L. Qiao et.al. [6] presented a watermarking scheme for MPEG Encoded Video and proved its non-invertibility by using an encryption function in the construction of the watermark to achieve the non-invertibility. The watermark construction requirements, property of the Self-Proof Class and the usage of different watermarking schemes had been discussed. Due to high computation penalty in DWT which is not suitable for real-time applications, DCT is chosen in this work to accomplish video copyright protection for real-time applications.

4 Video Compression

One of the most preferred techniques for compressing video is inter-frame compression. Inter-frame compression uses one or more earlier or later frames in a sequence to compress the current frame, while intra-frame compression uses only the current frame, which is effectively image compression. Inter-frame compression works by comparing each frame in the video with the previous one. If the frame contains areas where nothing has moved, the system copies that part of the previous frame, bit-by-bit, into the next one. Inter-frame compression works well for programs that will simply be played back by the viewer.

Video inter-frame compression is done using motion compensation and Discrete Cosine Transform (DCT) techniques. The block matching algorithm divides the current frame and the previous frame into several macro blocks, comparing the blocks in the two frames and trying to search for the best matched pairs for each block. Matching criteria is Sum of Absolute Differences (SAD), which is defined as follows:

$$O(j, k) = \sum_{m=0}^{(M_t-1)} \sum_{n=0}^{(N_t-1)} \text{abs}(I(m+j, n+k) - T(m, n)) \quad (1)$$

Where $I(m, n)$ is the input macro block and
 $T(M_t, N_t)$ is the template block values

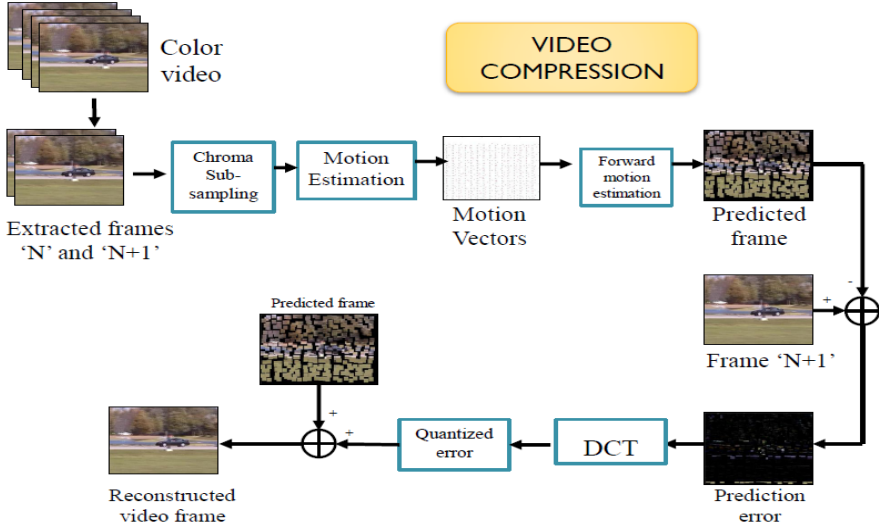


Fig. 2. Inter frame Video compression

The encoder calculates motion vectors between successive frames and uses them to reduce redundant information. The coder subtracts the motion-compensated prediction from the source picture to form a prediction error picture. The prediction error is transformed with the DCT, the coefficients are quantized, and these quantized values are coded. The coded prediction error is combined with other information such as motion vectors and synchronizing information, and formed into a bit stream for transmission. In the decoder, the quantized DCT coefficients are reconstructed and inversely transformed to produce the prediction error. This is added to the motion-compensated prediction generated from previously decoded pictures to produce the decoded output as shown in Fig. 2.

5 Compressed Domain Watermarking Algorithm

The reason for choosing compressed domain is due to

- Fast computation
- Consumes less memory
- Suits for Real-Time Applications

If $C(i, j)$ and $W(I, j)$ are host video frame and watermark image DCT coefficients respectively then watermarked video frame DCT coefficients are obtained by

$$C_w(i, j) = (\beta_n \times C(i, j)) + W(i, j) \tag{2}$$

Where β_n is the embedding factor, which ranges from 0 to 1. Higher the value of β_n , better the watermark perceptibility in the original video. A monochrome watermark image is embedded into Y color space only. The proposed watermarking algorithm is presented as a flow chart in fig. 3.

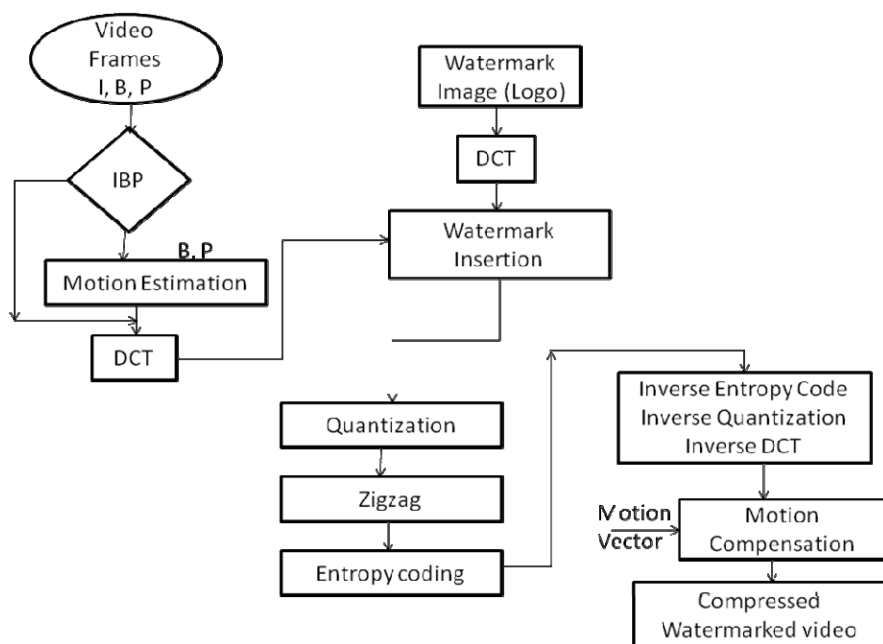


Fig. 3. Algorithm flow for compressed domain video watermarking

- Quantization module**
 Quantize the DCT coefficient according to standard quantization tables which is defined previously. The input and output are buffered to the frame buffer and inverse quantization is done to resume the original 8×8 DCT coefficient matrix.
- Zigzag module**
 Zigzag scanning of the DCT is Performed for re-ordering of the DCT coefficients. Inverse zigzag scanning is applied to resume the original order of 8×8 DCT coefficient matrix.
- Entropy coding module**
 Huffman coding look up is used to perform entropy. Inverse entropy coding is applied to the Huffman pre-calculated as decoding lookup table.

6 SIMULINK Prototyping

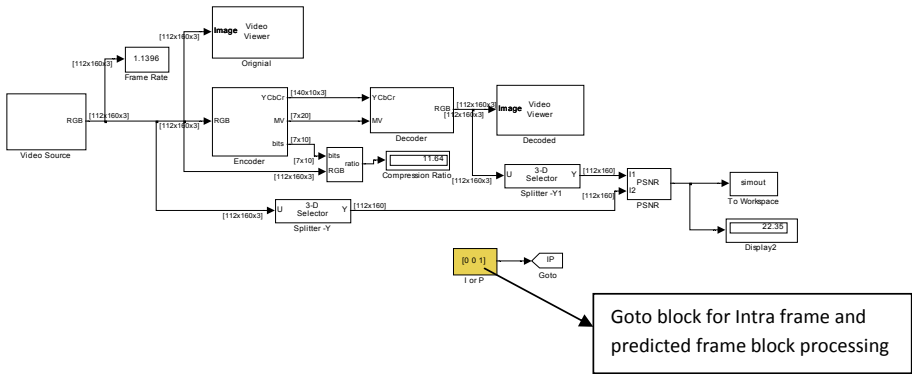
Simulink modeling of the individual architecture modules and the overall system is discussed here.

6.1 System Level Modeling

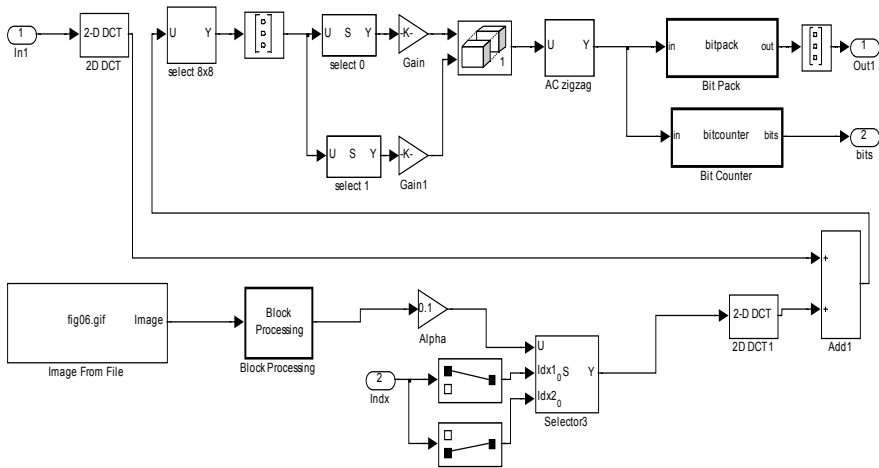
Using video and image processing block sets in Simulink, a fast prototyping module is built to verify algorithm and architecture. Simulink offers common video and image processing functions and modules. The available modules are DCT/IDCT, SAD for Motion Estimation, Block Processing and Delay (Buffer). Quantization, Zigzag scanning and Entropy coding blocks are built.

6.2 Modeling the Architecture

The Simulink system block set diagram for video watermarking in the compressed domain is shown in **Fig.4**. **Fig.4 (a)** presents the overall architecture of compressed domain video watermarking. The watermarking block in **Fig.4 (b)** embeds the watermark in all I, P frames.



(a) Overall Architecture



(b) The Watermarking Encoder YUV

Fig. 4. Simulink design for compressed domain video watermarking

7 Experimental Results

Exhaustive simulations to make assessment of watermarking quality with a large variety of watermark images and video clips are made. Performance of the watermarked clips and architecture is analyzed using various parameters.

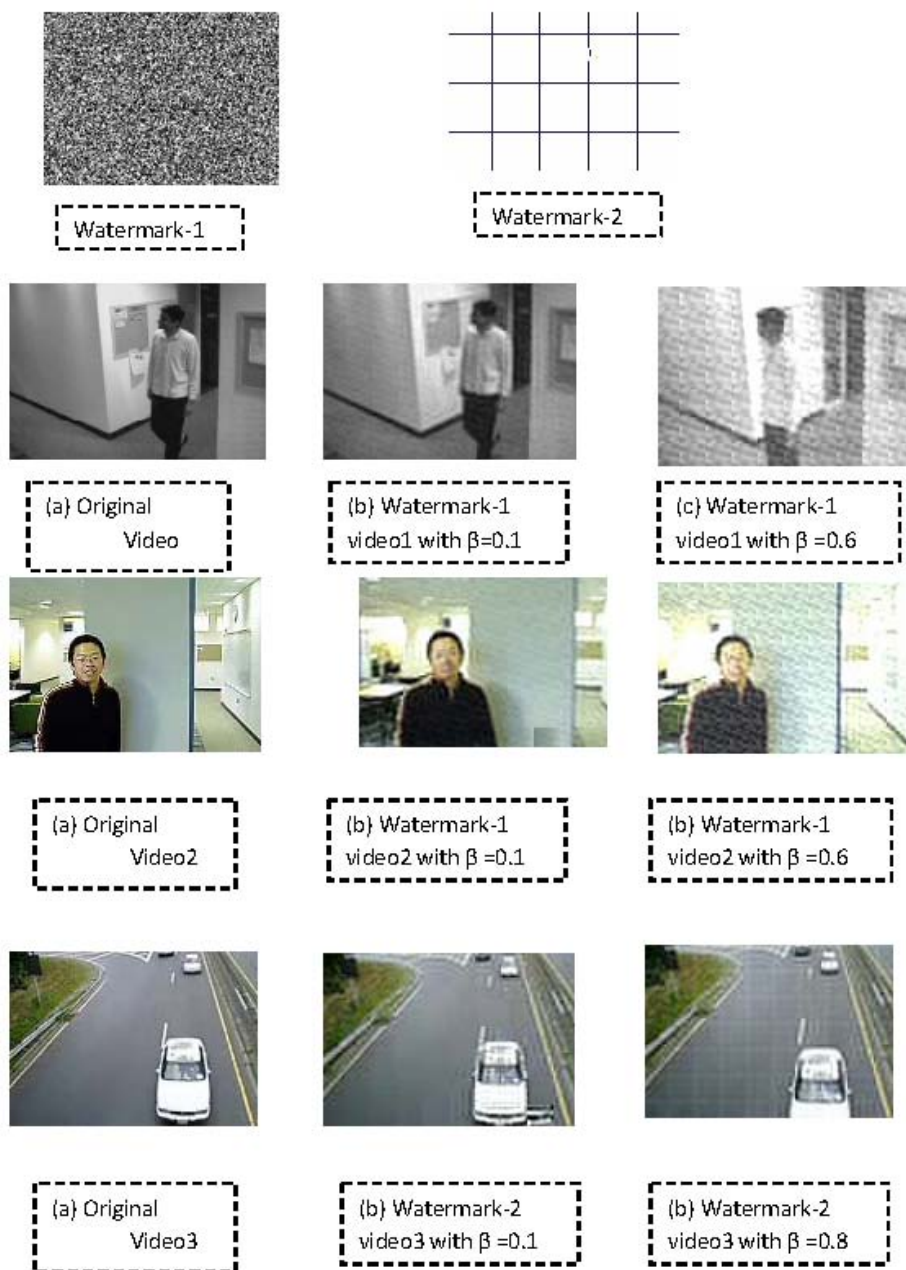


Fig. 5. Selected sample videos used in the experiments.

A sample watermark and watermarked video used in the experiment are presented in Fig.5. The following are the equations used to test the quality metrics of a video.

$$MSE = \frac{\sum_{M \times N} [I_1(m,n) - I_2(m,n)]^2}{M \times N} \tag{4}$$

$$PSNR = 10 \log_{10} \left(\frac{R^2}{MSE} \right) \tag{5}$$

Where M and N are the dimensions of input video frames and R is the maximum fluctuation in the input. If input image has an 8-bit unsigned integer data type, R is 255. From the above PSNR equation, the quality metrics of video compression and watermarking with different strength factor are given in Table I. The average PSNR of watermarked compressed video is (24-28) dB.

Table 1. Video quality metrics for Watermark 1

Clips	Scaling Factor α=0.1		Scaling Factor α=0.6	
	Compression Ratio (Average)	PSNR (in dB)	Compression Ratio (Average)	PSNR (in dB)
Video1 (46 Frames)	12.01	21.28	11.39	12.72
Video2 (96 Frames)	11.65	22.46	10.37	11.89
Video3 (46 Frames)	12.22	22.98	11.62	11.29

Table 2. Video quality metrics for Watermark 2

Clips	Scaling Factor α=0.1		Scaling Factor α=0.8	
	Compression Ratio (Average)	PSNR (in dB)	Compression Ratio (Average)	PSNR (in dB)
Video1 (46 Frames)	12.11	22.86	11.95	22.32
Video2 (96 Frames)	11.85	28.86	11.63	26.32
Video3 (46 Frames)	12.80	20.84	12.59	21.06

When selected samples are watermarked, it can be analyzed that the noise in watermark2 is dominant than watermark1. So, PSNR of watermark1 is reduced when compared with the watermark2 even though embedding factor (β) is considered as 0.6.

To estimate the similarities between extracted and original watermark cross-correlation is used.

$$C(l, f) = \sum_{m=0}^{(M_a-1)} \sum_{n=0}^{(M_b-1)} A(m, n) \cdot \text{conj}(B(m+l, n+f)) \quad (6)$$

Where A, B are the original and extracted watermarks of size (Ma, Mb) respectively. The normalized cross correlation is used as a parameter to judge the robustness of the watermark. Proposed method is tested through the following attacks:

1. Filtering
2. Salt and Pepper Noise
3. Rotation
4. Gaussian Noise
5. Speckle Noise

A video sequence of 100 frames is used to test several attacks. The visual content varies greatly in the video sequence.

7.1 Filtering

The experimental results obtained by applying nonlinear (median) filter are shown in **Fig.6**. It is analyzed that the strength of the proposed algorithm depends on the size of the filtering window. The smaller window provides better performance.

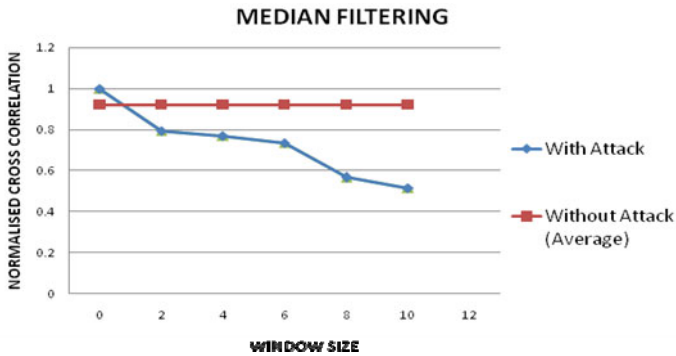


Fig. 6. Estimated normalized cross correlation value for different window size median filter operations

7.2 Salt and Pepper Noise

The experimental results are shown in **Fig.7**. This randomly occupies white and black (on and off) pixels.

7.3 Rotation

This falls under the category of geometric attacks, Table 2 shows the results at different angles of rotation. It is analyzed that the small amount of rotation is

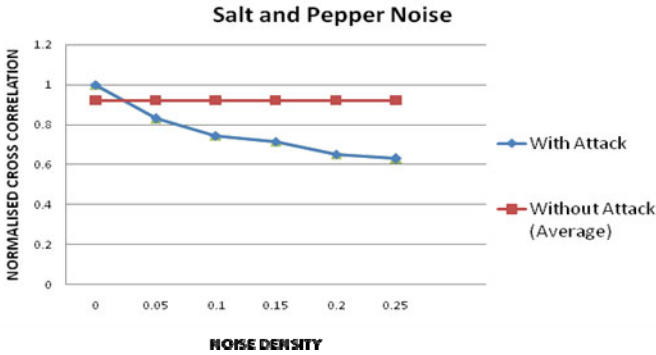


Fig. 7. Estimated normalized cross correlation value for different noise density values

acceptable. The rotation detector can increase the robustness with increased complexity.

Table 3. Results at different angles of rotation

Rotation Angle	Detection (%)
0	100
2	94
4	92
8	82

7.4 Gaussian Noise:

The Gaussian Noise block generates discrete-time white Gaussian noise. The output Gaussian random variables are uncorrelated. The experimental results are shown in Fig.8.

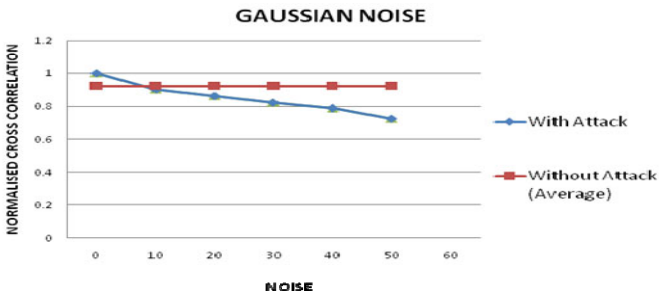


Fig. 8. Estimated normalized cross correlation value for different noise values

7.5 Speckle Noise

This attack adds multiplicative noise to image based on the following equation. Where 'n' is uniformly distributed random noise with mean 0 and variance v. The experimental results are as below

$$J = I + (n \times I) \quad (7)$$

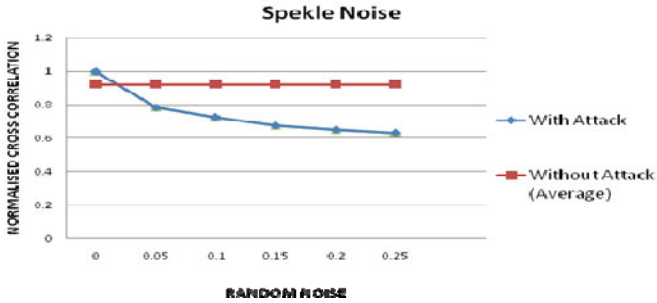


Fig. 9. Estimated normalized cross correlation value for different random noises

8 Conclusion

A frame work for video watermarking during compression which is robust to attacks is designed and tested. The proposed algorithm shows better NC which decides the robustness and security of the system. An average of about 8 % increase in Normalized Correlation is obtained for different video samples. The architecture for watermarking algorithm is prototyped using Simulink. Further research is under way to implement in field-programmable-gate-array (FPGA), to improve the PSNR.

References

1. Strycker, L.D., Termont, P., Vandewege, J., Haitsma, J., Kalker, A., Maes, M., Depovere, G.: Implementation of a Real-Time Digital Watermarking Process for Broadcast Monitoring on Trimedia VLIW Processor. In: IEEE Proceedings on Vision, Image and Signal Processing, vol. 147(4), pp. 371–376 (2000)
2. Mathai, N.J., Kundur, D., Sheikholeslami, A.: Hardware Implementation Perspectives of Digital Video Watermarking Algorithms. IEEE Transactions on Signal Processing 51(4), 925–938 (2003)
3. Garimella, A., Satyanarayan, M.V.V., Kumar, R.S., Muruges, P.S., Niranjan, U.C.: VLSI Implementation of Online Digital Watermarking Techniques with Difference Encoding for the 8-bit Gray Scale Images. In: Proceedings of the International Conference on VLSI Design, pp. 283–288 (2003)
4. Fan, Y.C., Van, L.D., Huang, C.M., Tsao, H.W.: Hardware Efficient Architecture Design of Wavelet-based Adaptive Visible Watermarking. In: Proceedings of 9th IEEE International Symposium on Consumer Electronics, pp. 399–403 (2005)

5. Biswas, S., Das, S.R., Petriu, E.M.: An adaptive compressed MPEG-2 video watermarking scheme. *IEEE Transactions on Instrumentation and Measurement* 54(5), 1853–1861 (2005)
6. Qiao, L., Nahrstedt, K.: Watermarking Methods for MPEG Encoded Video: Towards Resolving Rightful Ownership. In: *Proceedings of the IEEE International Conference on Multimedia Computing and Systems*, pp. 276–285 (1998)
7. Maes, M., Kalker, T., Linnartz, P.M.G., Talstra, Depovere, G.F.G., Haitzma, J.: Digital Watermarking for DVD Video Copyright Protection. *IEEE Signal Processing Magazine* 17(5), 47–57 (2000)
8. Tsai, T.H., Wu, C.Y.: An Implementation of Configurable Digital Watermarking Systems in MPEG Video Encoder. In: *Proceedings of the IEEE International Conference on Consumer Electronics*, pp. 216–217 (2003)
9. Xvid Codec, <http://www.xvid.org>

Decimal Floating Point Multiplier Using Double Digit Decimal Multiplication

Rekha K. James¹, K. Poulose Jacob¹, and Sreela Sasi²

¹ Cochin University of Science and Technology, Kochi, Kerala, India

² Gannon University, Erie, PA, USA

{rekhajames, kpj}@cusat.ac.in, sasi001@gannon.edu

Abstract. Floating-point representation can support a wider range of values over fixed point representation. The performance of Decimal Floating Point (DFP) operations plays an important role in financial and commercial computations. This paper uses an iterative decimal floating point multiplier using Double Digit Decimal Multiplication (DDDM) technique for decimal fixed point multiplication. It performs two digit multiplications simultaneously in one cycle. The floating point multiplier incorporates exponent processing, rounding and exception detection capabilities. The intermediate exponent, product sign, sticky bit, round digit and guard digit are determined on the fly during the accumulation of partial products. Simulation is done for 32-bit DFP data using the proposed approach. The results are compared and tabulated for area and delay with the existing design in literature.

Keywords: Decimal Multiplier, Floating Point, Double Digit Implementation, Rounding Logic, Low Latency.

1 Introduction

The majority of the world's commercial and financial data are stored and manipulated in decimal form. Currently, general purpose computers do decimal computations using binary arithmetic. Binary data can be stored efficiently and manipulated very quickly on two-state computers. However, there are compelling reasons to consider decimal arithmetic, particularly for business computations. The reasons include human's natural affinity for decimal arithmetic, and the inexact mapping between some decimal and binary values. Binary floating-point values can only approximate certain common decimal numbers. For example a value of 0.1 requires an infinitely recurring binary pattern of zeros and ones. When an average user performs decimal addition of 0.1 and 0.9, the result is 1.0. If the decimal addition is performed in binary, the result may be 0.99, due to error generated by the decimal to binary conversion. In this world of precision, such errors are no more tolerable. In many cases, the law requires that results generated from financial calculations performed on a computer should exactly match with those carried out using pencil and paper. This is possible only if the calculations are done in decimal. Recently, support for decimal arithmetic has received increased attention due to the growing importance in financial

analysis, banking, tax calculation, currency conversion, insurance, telephone billing and accounting. Hardware support for decimal operations, however, has been limited. The scenario is set to change with the continually dropping cost, and with the significant speedup achievable in hardware. This leads to the design of processors that will support Decimal Floating Point (DFP) arithmetic in near future.

Several hardware designs using IEEE 754–2008 standard [1] for DFP multiplication have been proposed in literature. The DFP multiplier in [2] makes use of decimal carry save adders [3] for Decimal Fixed Point (DFxP) multiplication of its significand digits. The design of DFP multipliers whose partial product accumulation is based on non-pipelined iterative implementation using decimal carry save adders is presented in [4]. A combined decimal and binary floating-point multiplier is presented in [5]. To attain high speeds, parallel multipliers are used at the expense of area. A parallel DFP multiplier using parallel DFxP multiplier is presented in [6, 7]. Parallel designs are adopted when latency and throughput are considered more important than area.

In most of the iterative DFP multipliers published so far, the entire multiplicand is multiplied by one multiplier digit to generate a partial product in each cycle. This paper uses a design for double digit DFxP multiplication using Binary Coded Decimal (BCD) encoding of [8]. It gives a DFxP multiplication algorithm suitable for high-performance with less number of cycles. This DFxP multiplier performs 2 digit multiplications simultaneously in one cycle. When multiplying two n -digit operands to produce a $2n$ -digit product, the DFxP design has latency of $\lceil (n/2) + 1 \rceil$ cycles. The approach for DFP multiplier design presented in this research utilizes double digit decimal multiplication (DDDM) for the DFxP multiplication.

The organization of the paper is as follows: Initially, the background information on IEEE 754–2008 DFP formats are presented. This is followed by the descriptions of intermediate exponent, rounding and exception handling. The approach for DFP multiplier design using the DDDM technique is then described. DDDM for 7 digits are synthesized with Leonardo Spectrum from Mentor Graphics Corporation using ASIC Library, and the results are tabulated. The area and delay factors of simulation results are compared with single digit implementation.

The DFP design also incorporates the necessary DFP exponent processing, rounding and exception detection capability. The intermediate exponent, product sign, sticky bit, round digit and guard digit are determined in parallel with the generation and accumulation of partial products. Simulation is done for 32-bit DFP data using DDDM for DFxP multiplication, and synthesized using Leonardo Spectrum from Mentor Graphics using ASIC Library. This result is then compared for area and delay with the existing design of [4].

2 DFP Formats

The IEEE 754–2008 standard specifies DFP formats of 3 representations: 32-bit format with 7 significand digits, 64-bit format with 16 significand digits and 128-bit format with 34 significand digits. These encodings allow a range of positive and negative values,

2.3 Round Digit and Guard Digit Generation

Rounding has to be done on the most significant ‘n’ digits based on the rounding schemes tabulated in Table1. When the final product of the DFxP multiplication is one digit less than the total length, it may be necessary to shift left the product by one digit to make the MSD a non-zero number. The corrective left shift of one digit necessitates maintaining an additional digit to the right of the decimal point. This digit is referred to as the guard digit and is similar to the guard bit used in binary floating point multiplication. The corrective left shift may lead to an overflow to the (n+1)th digit while rounding the result. This situation is demonstrated below using a 7 digit × 7 digit DFxP multiplier of a 32-bit DFP multiplication.

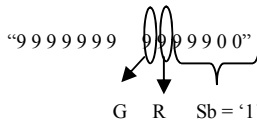
Table 1. Rounding Schemes

Rounding mode	Condition
Round up	G > 5 G = 5 and R ≠ 0 G = 5 and Sb ≠ 0
Round down	G < 5
Round to nearest even	G = 5 and R = 0 and Sb = 0

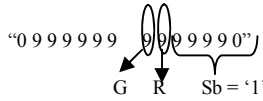
Let C1 be the significand of operand 1, and C2 the significand of operand 2.

If C1 = “3 333 330” and
 C2 = “3 000 003”
 Then the 14 digit result of C1 × C2 =
 “0 9 9 9 9 9 9 9 9 9 9 9 9 0”

Since the MSD is zero a left shift is performed to avoid leading zero and the exponent is adjusted accordingly.
 Now the new result is



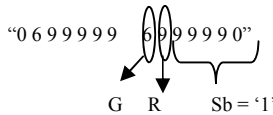
Since G > 5, the first 7 digits are to be rounded up, and results in an 8 digit number “10 000 000”. This overflow to the eighth digit has to be corrected again. This additional correction or shift can be overcome by doing the shift after rounding. So, in this case the rounding is done initially for the product and the result is shown below.



Even now, since $G > 5$, the most significant 7 digits are rounded up to get a value “1 000 000”. The MSD of the rounded result is not zero anymore and hence left shift is not required. Now consider another example.

If $C1 = “2 3 3 3 3 3 0”$
 and $C2 = “3 0 0 0 0 3”$

The 14 digit result of $C1 \times C2 =$



Since $G > 5$, the first 7 digits are rounded up, and results in a number “070000000”. The MSD of the rounded result is a zero, and hence a left shift is done and the exponent value is adjusted. The shifted result is now “7 0 0 0 0 0” if the ‘shifted in digit’ is zero, or a more erroneous result, “7 0 0 0 0 6” if the ‘shifted in digit’ is the Guard digit (G). Shifting before rounding would have given a result “6 9 9 9 9 7”. This result is equivalent to rounding the ‘n’ digits excluding the MSD. It also utilizes the correct purpose of Guard digit. It can be seen that by rounding the result before shifting the error generated is more. So, in such cases shifting has to be done first. In other words, select the result after rounding the ‘n’ digits excluding the MSD. A suitable selection method is required to determine if rounding is to be done before shifting or vice versa.

3 DFP Multiplication

The approach for DFP multiplier design presented in this paper makes use of the iterative DFxP multiplier design using Double digit Decimal Multiplication (DDDM).

3.1 Double Digit Decimal Multiplication (DDDM)

The block diagram for the DDDM proposed in [8] is shown in Fig. 2. The DFxP multiplier performs 2 digit decimal multiplications simultaneously in one cycle. Suitable secondary multiples are selected by using two pairs of multiplexers for the two digit multiplier shift register output using Table 2. For each iteration cycle the multiplicand is multiplied by 2 digits of the multiplier.

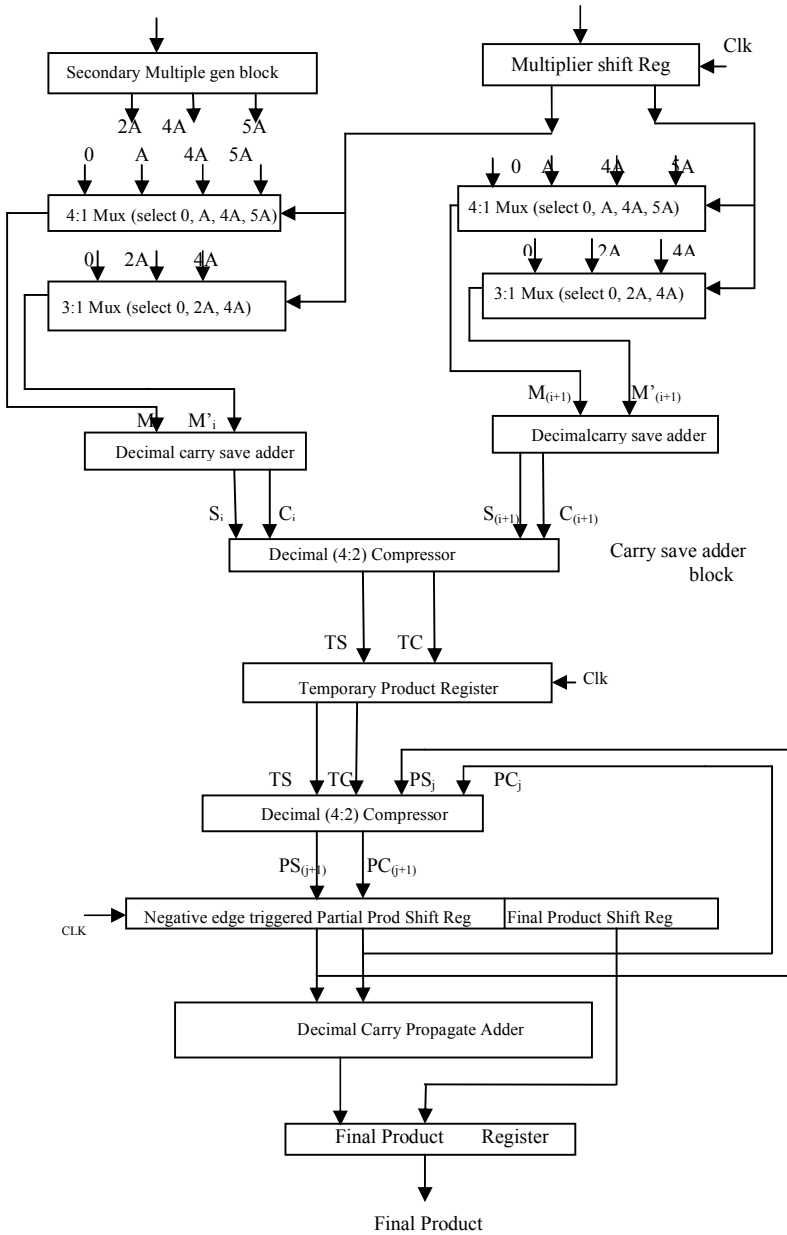


Fig. 2. Block Diagram of the Double Digit Fixed Point Decimal Multiplication

Table 2. Recoding of Digits of Bi

B _i	M _i	M' _i	B _i	M _i	M' _i
0	0	0	5	5A	0
1	A	0	6	4A	2A
2	0	2A	7	5A	2A
3	A	2A	8	4A	4A
4	4A	0	9	5A	4A

The partial product formed is shifted by 2 digits and the process is repeated for $\lceil (n/2) \rceil$ iterations. After $\lceil (n/2) \rceil$ iterations the final product in the form of ‘carry save’ sum and carry is available at the output of the ‘Partial Product Shift Register’. This is then passed to a ‘Decimal Carry Propagate Adder’ (CPA). The adder output is then stored in the ‘Final Product Register’. The final product is available after $\lceil (n/2) + 1 \rceil$ cycles.

3.2 DFP Multiplication Using DDDM

The design for DFP multiplication is shown in Fig. 3. The operands encoded in DPD are decoded to get BCD significand digits and binary exponents. Then the product of significand digits of the two operands is generated using DDDM of iterative DFxP multiplication. The Sticky bit (Sb), Round digit (R) and Guard digit (G) generation are done in parallel with this process. Rounding is accomplished by selecting either the product truncated to ‘Plimit’ digits or its incremented value. Initiation of the rounding operation along with the DFxP multiplication speeds up the entire process. After rounding, a ‘zero’ at MSD leads to a single digit shift towards left. The shifter module is a major component that contributes to the critical path delay of a floating point multiplier. So, a second rounding module is used to round the product, excluding the MSD. The rounded result is selected from these two results based on MSD. The exponent is adjusted, exceptions are set accordingly, and the result is encoded back in DPD.

The algorithm is explained using a 32-bit DFP multiplication. Initially the two 32-bit operands are read from registers, and decoded to generate 7 significand digits, 8-bit biased exponent (E) and a Sign bit (S). The exceptions such as ‘NaN’ and ‘Infinity’ are also decoded out from the 32-bit input. The output exceptions are set at the logic block named ‘Handle exception 1’ depending on the input exceptions.

There are four exceptions that may be signaled during multiplication: Invalid operation, overflow, underflow, and inexact. The exponent is computed, and the sign bit of the result is determined as the XOR combination of sign bit of the two operands. The significand digits are multiplied using DDDM based DFxP multiplier. A 7 digit \times 7 digit DFxP multiplier is used for 32-bit DFP input. The least significant ‘n’ digits of the ‘2n’ digit DFxP product are available after $\lceil n/2 \rceil$ cycles. The most significant ‘n’ digits are generated in $\lceil (n/2) + 1 \rceil^{th}$ cycle by the ‘Decimal Carry Propagate adder’ of the DFxP unit. The Sticky bit (Sb), Round digit (R) and Guard

digit (G) generation depends on the least significant ‘n’ digits and so are generated in $\lceil (n/2) + 1 \rceil^{th}$ cycle. Rounding is done in the $\lceil (n/2) + 2 \rceil^{th}$ cycle and is accomplished by 2 rounding modules as explained earlier.

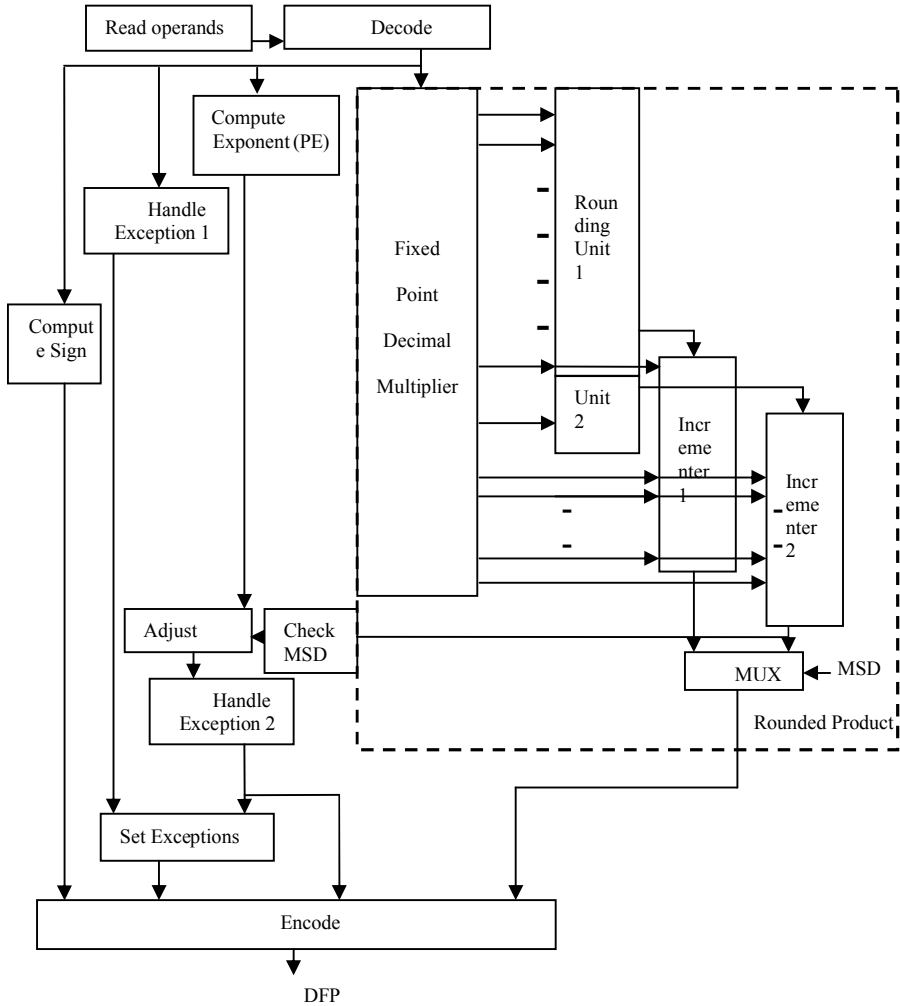


Fig. 3. Decimal Floating point (DFP) Multiplier

The MSD of the significand of each input operand is suggested to be a non-zero number. This leads to a unique representation of the DFP inputs avoiding the need to count the number of leading zeroes. The smallest number that can be represented in this format for a 32 bit representation is 1000000×10^{-101} and the largest number is $9999999 \times 10^{+90}$. Any number that is greater than the largest number is encoded as

'Infinity' with an 'Overflow' exception. Similarly any number that is less than the smallest number is truncated to 'Zero' with an 'Underflow' exception. If the result from an operation needs more digits than 'Plimit' which is the maximum precision of the significand, then the result will be rounded. If this rounding causes removal of non-zero digits then, the 'Inexact' exception is set.

In the proposed DFP multiplier, rounding is done for both the most significant 'n' digits (using Incrementer 2 of Fig. 3) and for the 'n' digits excluding MSD (using Incrementer 1 of Fig. 3). The appropriate result is selected based on the MSD of the Incrementer 2. This improves the accuracy of the result if a left shift is required, and also avoids the need for a shifter module. This in turn reduces the critical path delay. After rounding, the required 'exponent adjust' is done and if necessary, exceptions are modified. The final result is then encoded back to DFP format.

4 Synthesis Results of DFP Multiplier

The proposed DFP multiplier using DDDM technique for DFxP multiplication is synthesized for a 32-bit input to find the area and delay associated with the design. Synthesis was done using Leonardo Spectrum from Mentor Graphics Corporation with ASIC Library. An area and delay breakdown for an approximate contribution of major components of the design shown in Fig. 3 is given in Table 3.

Table 3. Area and Delay for different stages of DFP Multiplier (32-bit)

Component	Area	Delay
	μm^2	ns
Decoding Logic	1439	3.26
Exponent generation	229	4.42
Exception handling	66	4.21
DFxP multiplier & Rounding Unit	20105	41.1
Encoding Logic	600	5.25
Total	22439	52.97

Even though the total area is the sum of area of different stages, the total delay for the complete circuit is only 52.97ns, which is less than the total sum of delays of all stages. This is because of the inherent parallelism in the design. Table 4 shows the synthesis results of DFxP Multiplier and rounding unit of Fig. 3. Here, also the total delay is less than the sum of delays of each component as rounding starts before the DFxP multiplication is completed. This is possible since the sticky bit (Sb), the round digit (R) and the guard digit (G) are generated on-the-fly during the DFxP multiplication process. For a 7 digit \times 7 digit DFxP multiplication, the sticky bit (Sb) generation can be done after the fourth cycle; round and guard digit generation is done in the next 2 cycles. The delay break up of different components of DFP Multiplication is shown in Fig. 4.

Table 4. Area and Delay of Rounding Unit and DFxP Multiplier (7 digit × 7 digit)

Component	Area	Delay
	μm ²	ns
DFxP Multiplier	12713	18.16
Incrementer 1	3536	22.71
Incrementer 2	3536	22.71
Mux	205	0.24
Sticky bit	51	1.15
Rounding	64	1.96
Total	20105	41.1

When multiplying two DFP numbers with ‘n’ digit significands, using this proposed approach, the worst case latency is $\lceil (n/2) + 2 \rceil$ cycles, and initiation interval is $\lceil (n/2) + 1 \rceil$ cycles. In other words, a new multiplication can begin every $\lceil (n/2) + 1 \rceil^{th}$ cycle. The final cycle of current set of inputs and the first cycle for next set of inputs are done simultaneously.

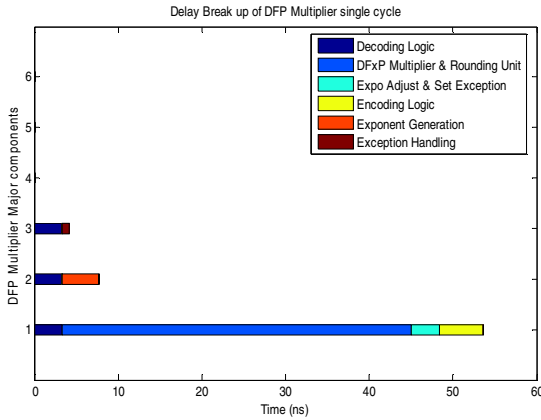


Fig. 4. Delay Break up of DFP Multiplier

DFP multiplier takes 6 cycles to complete one 32-bit DFP multiplication with worst cycle time being 31.56 ns. Hence, the total delay is 189.36 ns. The proposed design using DDDM requires lesser number of clock cycles for DFP multiplication compared to the DFP multiplication in [4]. This in turn reduces the total delay by 21.10% even though the worst case cycle time is more. Both these designs are synthesized in the same environment. The comparisons of delay of the proposed approach for DFP multiplication for a 32-bit input with the design of [4] is given in

Table 5. Extending the iterative DFxP multiplier design to support DFP multiplication affects the area, cycle time, latency, and initiation interval.

Table 5. Comparison of DFP Multipliers for 32-bit input

Parameters	DDDM	Design of [4]
Worst cycle time (ns)	31.56	26.67
Maximum frequency (MHz)	31.68	37.5
Delay of 32-bit DFP multiplication in terms of worst cycle time (ns)	189.36	240.03

5 Conclusion

The iterative DFP multiplier presented in this research includes the floating point extensions to the iterative DFxP multiplier design using DDDM. This DFP multiplier design is in compliance with IEEE 754-2008 standard. A 32-bit DFP multiplier is synthesized to find the area and delay using Leonardo Spectrum from Mentor Graphics Corporation with ASIC Library. The proposed design and the design in [4] for the DFP Multiplication are synthesized in the same environment. The DFP design using DDDM for DFxP multiplication requires lesser number of clock cycles for DFP multiplication compared to that in [4]. The latency to complete ‘n’ digit × ‘n’ digit DFxP multiplication is almost halved compared to single digit design with an increase in area of 50%. The latency for the multiplication of DFP numbers with ‘n’ digit significands using DDDM is $\lceil (n/2) + 2 \rceil$ cycles, and a new multiplication can begin every $\lceil (n/2) + 1 \rceil$ cycle. This in turn reduces the total delay, even though the worst case cycle time is more, giving a delay reduction of 21.10% in comparison to the design in [4].

References

1. IEEE Working Group of the Microprocessor Standards Subcommittee, IEEE Standard for Floating-Point Arithmetic. The Institute of Electrical and Electronics Engineers, New York (2008)
2. Erle, M.A., Schulte, M.J., Hickmann, B.J.: Decimal Floating-Point Multiplication Via Carry-Save Addition. In: 18th IEEE Symposium on Computer Arithmetic, pp. 46–55. IEEE Computer Society, Los Alamitos (2007)
3. Erle, M.A., Schulte, M.J.: Decimal Multiplication via Carry-Save Addition. In: 14th IEEE International Conference on Application-Specific Systems, Architectures, and Processors, pp. 348–358 (June 2003)
4. Erle, M.A., Hickmann, B.J., Schulte, M.J.: Decimal Floating-Point Multiplication. IEEE Transactions on Computers 58(7), 902–916 (2009) ISSN:0018-9340

5. Tsen, C., González-Navarro, S., Schulte, M., Hickmann, B., Compton, K.: A Combined Decimal and Binary Floating-Point Multiplier. In: 20th IEEE International Conference on Application-Specific Systems, Architectures and Processors, pp. 8–15 (2009)
6. Fahmy, H.A.H., Raafat, R., Abdel-Majeed, A.M., Samy, R., ElDeeb, T., Farouk, Y.: Energy and Delay Improvement via Decimal Floating Point Units. In: Proceedings on 19th IEEE Symposium on Computer Arithmetic, pp. 221–224 (2009)
7. Vazquez, A., Antelo, E., Moutuschi, P.: Improved Design of High-Performance Parallel Decimal Multipliers. *Proceedings of the IEEE Transactions on Computers* 59(5), 679–693 (2010)
8. James, R.K., Jacob, K.P., Sasi, S.: Double Digit Decimal Multiplier on Xilinx FPGA. In: Proceedings of the International Conference on Embedded System and Applications, ESA 2009, pp. 47–53 (2009)

Fuzzy Cost Based Multicast Routing for Mobile Ad-Hoc Networks with Improved QoS

G. Santhi¹ and Alamelu Nachiappan²

¹ Assistant Professor, Department of IT,
Pondicherry Engineering College
shanthikarthikeyan@pec.edu

² Associate Professor, Department of EEE,
Pondicherry Engineering College
nalam63@pec.edu

Abstract. In this paper we describe an idea of selecting best paths from source to destination node in Mobile Ad-hoc Networks (MANETs) using fuzzy cost. Here the multicast routing is carried out by selecting the most effective path in terms of minimum fuzzy cost by considering multiple independent QoS metrics such as bandwidth, end to end delay and number of nodes to transfer data from the source to the destination. In our method, the available resources of a path are converted into a single metric fuzzy cost. The fuzzy cost is calculated based on multi criterion objective fuzzy measure. The proposed fuzzy cost method is evaluated and compared with conventional protocol MAODV. Simulation results show that the proposed system performs better than the conventional MAODV in terms of improving packet delivery ratio and minimizing the end to end delay. The proposed multicasting protocol is simulated using the NS-2, while the fuzzy cost is calculated using Matlab.

Keywords: MANETs, Fuzzy Logic, Membership function, fuzzy cost, bandwidth, end to end delay, Quality of Service (QoS).

1 Introduction

Mobile Ad hoc Networks (MANETs) are a collection of mobile nodes that are dynamically and arbitrarily located in a way that interconnections between the nodes that are present in the network are capable of changing on a continual basis [1]. MANETs have been receiving a lot of importance for the last few years due to the rapid expansion of mobile devices and the ever increasing interest in mobile communications. MANETs consists of an unstructured network. The topology of any MANET is ever changing with time. The biggest challenge faced in MANET is routing. The primary goal of any ad hoc network routing protocol is to provide correct and efficient route establishment between pair of nodes so that messages may be correctly delivered in time. There exist numerous routing paths from source to destination node to transfer data; one of the routing paths is to be selected by any routing algorithm [11].

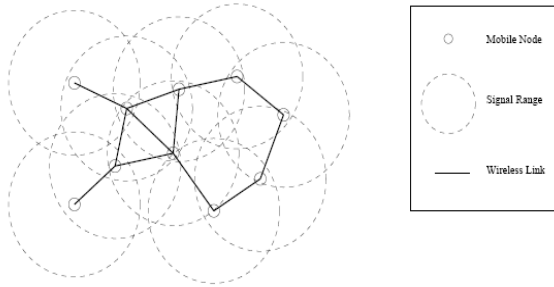


Fig. 1. MANET Environment

Our work mainly involves the selection of the most effective routing paths to transfer data. Hence the resources and traffic states of these routing paths should be monitored systematically to avoid creating traffic load or bottlenecks [12]. We concentrate on selecting appropriately effective path based on fuzzy cost instead of path selection by number of intermediate nodes.

This paper consists of the following sections. Section 2 describes about the Fuzzy Cost Based Multicast Routing and about the Fuzzy Cost Evaluation. Section 3 is about the design of proposed Routing Protocol. The route request/reply processes are also discussed in detail in this section. Section 4 is about the Simulation model and Section 5 discusses the conclusion.

2 Fuzzy Cost Based Multicast Routing

In this protocol we assume that n paths will exist from source to destination. The fuzzy cost of each path is based on the details of resources available in a path which include bandwidth, end to end delay and the number of nodes in between the source and the destination. Fuzzy cost is unique for each path, because we are applying linearity instead of nonlinearity [2].

2.1 Fuzzy Concept in MANETs

However, selecting a route which satisfies all multiple constraints is an NP complete problem [7]. There is no accurate mathematical model to describe it. Fuzzy logic is used to provide a feasible tool to solve the multimetric QoS problem. Fuzzy logic is a theory that not only supports several inputs, but also exploits the pervasive imprecision information [8]. So adopting fuzzy logic to solve multi metric problems in ad hoc networks is an appropriate choice.

In the literature very few routing algorithm will exists for MANET using fuzzy logic. These are Fuzzy Logic Wireless Multipath Routing (FLWMR) [9] and Fuzzy Logic Load Aware Multipath Routing (FLWLAR) [9]. The routing algorithm FLWMR considered only the metric is hop count and in FLWLAR the QoS metric is traffic load

along the link are input to the fuzzy controller, based on these metrics fuzzy controller evaluates the fuzzy cost. But the proposed algorithm considers five characteristics of network to find the fuzzy cost. In FLWMR and FLWLR fuzzy controller was designed based on nonlinear property where as the proposed method introduced linearity when evaluating the fuzzy cost.

2.2 Fuzzy Cost Evaluation

In fuzzy cost evaluation, we consider k number of resources, of which some are favorable for routing and some are not favorable for routing. For example, the maximum bandwidth available is favorable for routing and maximum delay is not favorable for routing.

Several metrics have been chosen to meet these objectives and to produce a single cost metric for selecting routes. The various routing metrics used are: bandwidth, delay and the number of nodes between the source and destination.

$$C = f(BA, D, N) \quad (1)$$

where,

BA – Available Bandwidth

D – End to end delay

N – Number of intermediate nodes between the source and destination

The bandwidth of a path is taken as the minimum available bandwidth of a node along the path.

$$BA(p(s,d)) = \min \{ B(e) \mid e \in P(s,d) \} \quad (2)$$

End to end delay of a path is the summation of the link delay at each link on the path [10].

$$D(p(s,d)) = \sum_{e \in P(s,d)} D(e) + \sum_{n \in P(s,d)} D(n) \quad (3)$$

where $P(s,d)$ is the path from the source 's' to the destination 'd' and 'e' is any link of the multicast tree.

2.3 Design of Fuzzy Inference Engine for FC-MAODV

A fuzzy engine shown in the figure below is used to calculate fuzzy cost by fuzzy logic and is typified by the inference system that includes the system rule base, input membership functions that fuzzify the input variables and the output variable is obtained by the de-fuzzification process. Fuzzification is a procedure where crisp input values are represented in terms of the membership function, of the fuzzy sets. The fuzzy logic controller triangular membership functions are defined over the range of the fuzzy input values. Following the fuzzification process the inference engine determines the fuzzy output using fuzzy rules that are in the form of IF - THEN rules. De-fuzzification is then used to translate the fuzzy output to a crisp value.

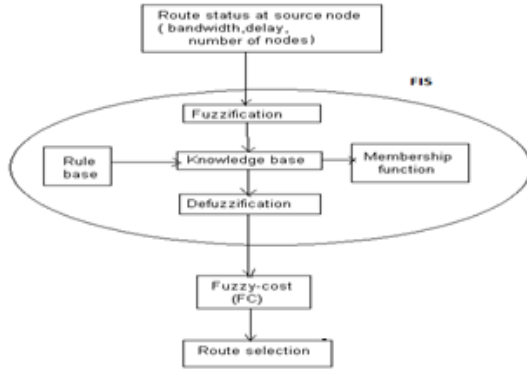


Fig. 2. Selection of multi-objective optimal route

Step 1. Fuzzification of Inputs and Outputs:

The three input variables to be fuzzified are available bandwidth, end to end delay and the number of intermediate nodes. On the existing knowledge of MANET, the metric bandwidth is represented by the variables *low*, *medium* and *high*. The end-to-end delay is represented by the variables *short*, *medium* and *long*. The number of nodes is represented by the variables *short*, *medium* and *high*. The output variable cost is represented by the variables *low*, *medium* and *high*. Each input (routing metric) is first fuzzified using fuzzy set manipulation. Triangular membership functions are used for fuzzification and have been extensively used for real time operations since they provide simple formulas and computational efficiency.

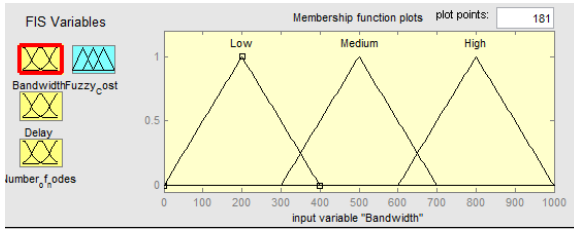


Fig. 3.a. Membership Function for Bandwidth

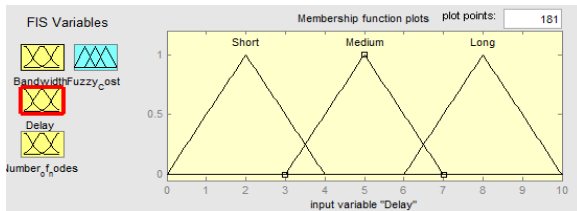


Fig. 3.b. Membership Function for Delay

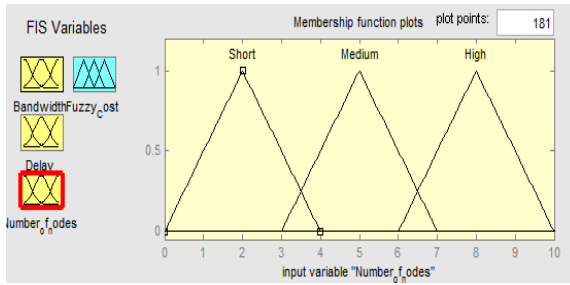


Fig. 3.c. Membership Function for Number of nodes

Step 2. Knowledge Base Rule Structure

The knowledge base (KB) is a set of rules (developed using expert knowledge) required to achieve the objective. The rules connecting the inputs and the output are designed based on a thorough understanding of the system. The fuzzy rules have IF-THEN structure. The inputs are combined using the AND operator. An example of rules which describes the input output mapping is as follows:

If (Bandwidth is High) and (Delay is short) and (Number_of_nodes is Short) then (Fuzzy_Cost is low)

Since the maximum bandwidth, minimum delay and less number of nodes are favorable for routing, here the fuzzy cost is low. If the path has minimum bandwidth, maximum delay and high number of intermediate nodes, then the fuzzy cost will be maximum, because they are unfavorable metrics for routing.

Step 3. Defuzzification

Defuzzification refers to the way a crisp value is extracted from a fuzzy set as a representation value. There are many kinds of defuzzifiers, here we take the centroid of area strategy for de-fuzzification. Since there are 3 inputs to Fuzzy Logic System and all the inputs are represented by three variables, there are 27 rules in total. Once the de-fuzzification process is done, the required fuzzy cost is calculated by the destination.

3 Proposed Routing Protocol

Every node in MANET acts as both a terminal and a router. Each node can become a destination for data traffic, thus, FLS is embedded in every mobile node. FCMAODV builds routes using a route request / route reply query cycle. When a source node desires a route to a destination for which it does not already have a route, it broadcasts a route request (RREQ) packet across the network. To calculate the fuzzy cost, the required node attributes of bandwidth and end to end delay fields are additionally included in the RREQ packet.

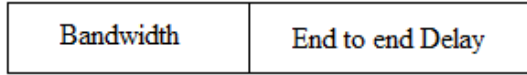


Fig. 4. New added fields to head of RREQ

Nodes receiving this packet update their information for the source node. If the source node finds the route in the cache, it directly sends the RREQ message and the route established and the transmission takes place. If the route is not in the cache, the source node broadcasts route requests across the network. The receiving node checks the threshold bandwidth (B_{th}) of the node and if it is satisfied, it updates its own information and broadcasts. If a node is the destination, the fuzzy cost is calculated for all the route requests that reach the destination. The route with minimum fuzzy cost is sent as route reply to the source via the intermediate nodes. The route discovery process is given in the following figure 5.

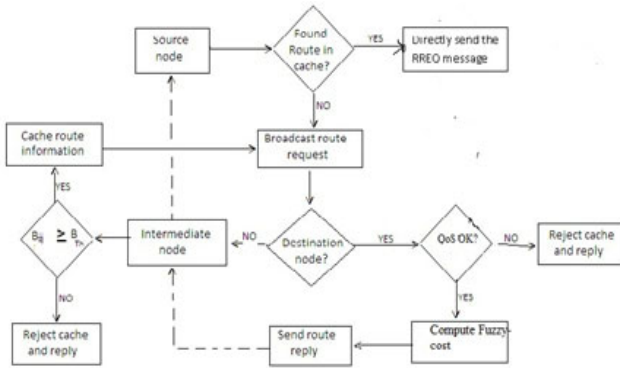


Fig. 5. Route discovery process

4 Simulation Model

A simulation model has been developed to test the application of the fuzzy logic system in the selection of better quality routes. The performance of Fuzzy and non-Fuzzy based routing is evaluated with varying path cache capacity according to the following metrics:

- Packet Loss Ratio : fraction of packets sent by a source node that does arrive at the destination node.

$$\text{Packet Delivery Ratio} = 1 - \text{Packet Loss Ratio}$$

- Control Overhead: The total number of control packets (RREQ, RREP, RERR) received by the destination node.

- End to end Latency: Delay incurred by a packet being transmitted between a source and destination node.

4.1 Analysis of Results

End to End Delay:

The path with minimum fuzzy cost will have the minimum delay because the resources are reserved and the route failure is minimum. The route failures have an impact on the delay, because the route failures require recalculation of routing and storing of packets in the sent buffer. Hence the end to end delay experienced in FCMAODV is less than that experienced in MAODV.

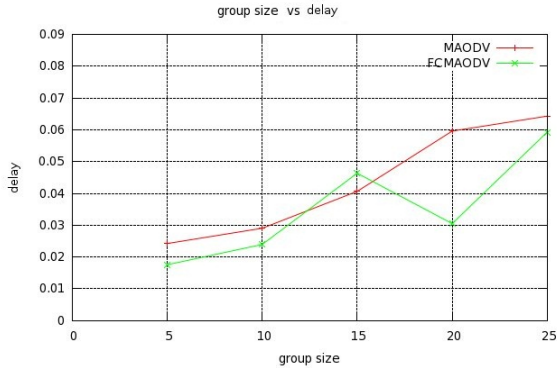


Fig. 6. Group size vs Delay

Packet Loss Ratio:

Packet loss ratio of the proposed system is less than MAODV for all group size values (Fig. 7.a). This is because of finding more reliable path and managing node breakage thereby avoiding the recomputation of route. As the group size increases the packet loss ratio of FCMAODV is less than that of MAODV which is due to its ability to select a set of stable and least congested routes thus having the lowest amount of congestion loss and very few route failures.

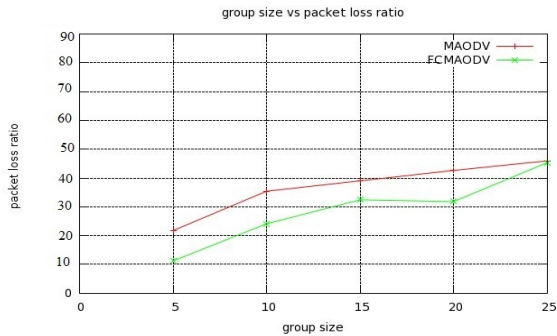


Fig. 7.a. Group Size vs Packet Loss Ratio

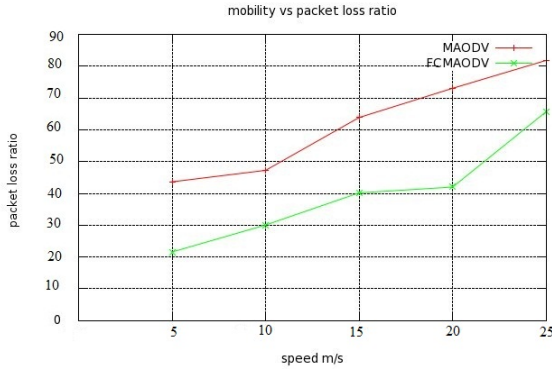


Fig. 7.b. Mobility vs Packet Loss Ratio

Fig 7.b) shows the performance of the average packet loss ratio under various mobility speeds. When the mobility is low, the multicast tree structure was mostly static and therefore the packet loss ratio is low. In FCMAODV, we used the multipath scheme to avoid the reconstruction of new tree when the links broke down due to mobility. Hence the packet loss ratio of FCMAODV is higher than that of MAODV.

Control Overhead:

Based on the minimum fuzzy cost, the route is decided by the destination itself and the route reply packet is sent through only one path with minimum fuzzy cost to the source. But, in MAODV, the route reply packets are sent from the destination to the source in all possible paths. Hence the Control Overhead of FC-MAODV is less when compared to of MAODV.

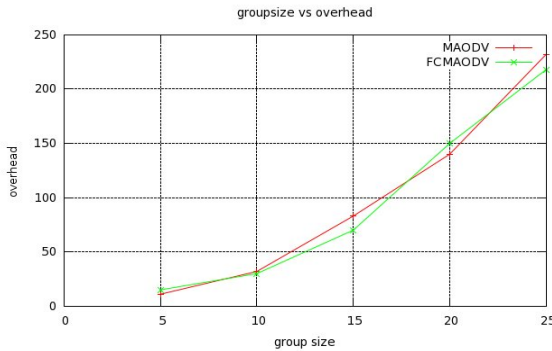


Fig. 8. Group Size vs Overhead

5 Conclusion

The proposed system considers QoS requirements for real time applications. Accounting for the uncertainty of route information in mobile ad-hoc network, fuzzy logic system is

adopted. It takes the different network state parameters such as bandwidth, end to end delay and number of intermediate nodes for calculating the fuzzy cost. Benefiting from all these optimizations, the improved protocol can be better performance than conventional MAODV protocol in terms of improving packet delivery ratio and minimizing the end to end delay.

References

1. Lian, J., Zhu, X.: A Multiple QoS Constraints Routing Protocol based on Mobile predicting in Ad Hoc Network. IEEE, Los Alamitos (2007)
2. Zuo, J., Chi, X., Lin, G., Li, H.: Service-aware Multi-constrained Routing Protocol With QoS Guarantee Based on Fuzzy Logic. IEEE, Los Alamitos (2008)
3. Manvaha, S., Srinivasan, D., Tham, C.K.: Evolutionary Fuzzy Multi-Objective Routing For Wireless Mobile Ad Hoc Networks
4. Thaw, M.M.: Fuzzy-based Multi-constrained Quality of Service Distance Vector Routing Protocol in Mobile Ad-hoc Networks. University of Computer Studies, UCSY Yangon, Myanmar
5. Naga Raju, A., Ramachandram, S.: Fuzzy Cost Based Multipath Routing For Mobile Ad-Hoc Networks. IEEE, Los Alamitos (2008)
6. Rea, S., Pesch, D.: Multi-Metric Routing Decisions for Ad Hoc Networks using Fuzzy Logic. IEEE, Los Alamitos (2008)
7. Alandjani, G., Johnson, E.E.: Fuzzy Routing in Ad Hoc Networks. In: 2003 IEEE International Conference on Performance, Computing, and Communications, April 9-11, pp. 525–530 (2003)
8. Zhang, X., Cheng, S., Feng, M.-y., Ding, W.: Fuzzy Logic QoS Dynamic Source Routing for Mobile Ad Hoc Networks. In: The Conference of Computer and Information Technology, CIT (2004)
9. Sun, B.L., Li, H., Zeng, Y., Qin, Q.Q.: Fuzzy QoS Controllers in Mobile Ad Hoc Networks using Genetic Algorithms. *Journal of Computational Information Systems* 3(6), 2255–2260 (2007)
10. Sun, B.L., Li, L.Y.: A Distributed QoS Multicast Routing Protocol in Ad Hoc Networks. *Journal of Systems Engineering and Electronics* 17(3), 692–698 (2006)
11. Luo, J., Xue, L., Ye, D.: Research on Multicast Routing Protocols for Mobile Ad-hoc networks. *Computer Networks* 52, 988–997 (2008)
12. Royer, E.M., Perkins, C.E.: Multicast Operation of the Ad hoc on demand distance vector routing protocol. In: Proc. ACM MOBICOM, pp. 207–218 (August 1999)

Cooperative Resource Allocation for Primary and Secondary Users with Adjustable Priorities in Cognitive Radio Networks

Behdis Eslamnour, S. Jagannathan, and Maciej Zawodniok

The Department of Electrical and Computer Engineering
Missouri University of Science and Technology
Rolla, MO 65409, USA

{ben88, sarangap, mjzx9c}@mst.edu

Abstract. The ability of cognitive radio to sense the spectrum and adapt its parameters in response to the dynamic environment makes it an ideal candidate for resource allocation of spectrum in wireless networks, especially co-existing and emergency networks. In the two latter networks, the secondary users should sense the spectrum and adapt their parameters such that they can use these resources without causing a degradation or interference to the performance of the primary/licensed users. Therefore, in this paper, a decentralized game theoretic approach for resource allocation of the primary and secondary users in a cognitive radio networks is proposed. In this work, the priorities of the networks are incorporated in the utility and potential functions which are in turn used for resource allocation. It is demonstrated analytically by using the potential and utility functions and through simulation studies that a unique NE exists for the combined game with primary users (PU) and secondary users (SU), and the combined game converges to the NE.

Keywords: cognitive radio networks, co-existing networks, resource allocation, game theory, Nash Equilibrium.

1 Introduction

Cognitive networks [1] have emerged as a new communication mode promising improved spectrum utilization and faster and more reliable network service. The two distinct characteristics of cognitive radios are *cognitivity* and *configurability* which enable the cognitive radios to sense the dynamic environment and adapt their parameters to optimize their performance.

Game theoretic approach is one of the ways to perform resource allocation, organize the networks and their users' behavior within the network and also with respect to the co-existing networks [2]-[8],[10],[13]. Existing game theoretic approaches such as [8] and [10] have taken a non-cooperative approach to power, joint power and rate control with Quality of Service (QoS) constraints, respectively. By contrast, a cooperative approach was taken in [3] to maximize the channel capacity. In [7], a framework for cooperation among the primary and secondary users is introduced. An interference avoidance

protocol for peer-to-peer networks was presented in [4] that assumed either the network is centralized, or the receivers are co-located.

The utility function used in the literature is a function of bit-error-rate [5],[8],[10], a combination of the signal power and interference power at the receiver [3], or signal-to-interference ratio [5]. Depending on the approach, the convergence and optimality of the algorithms are proven by using the existence of either Pareto optimality or Nash equilibrium (NE), existence of potential functions as well as the uniqueness of NE. However, the existing approaches [4],[5],[8] and [10] are applicable to a single network, or networks with equal priorities.

In [5], where primary and secondary networks are considered, it is assumed that there is only one primary user in the network. Though this work can be generalized to more than one primary user, the utility function is not discussed which takes into account the interference caused by the additional primary users (PUs). The work of [3] considers primary and secondary networks with the primary network using time division multiple access (TDMA) protocol where PUs cannot transmit at the same time. Unless the PUs communicate to a base station, the interference caused by a hidden node during communication among PUs, is not addressed.

On the other hand, authors in [5][7][10] have provided the presence of a unique NE. However, the convergence of the combined game with primary and secondary users is not shown since finding a utility function that provides a unique NE for the combined game is rather difficult. Others assumed a single game with homogeneous players [4], or only consider a single PU node and a set of SU players [3],[5]. By contrast, in this paper a decentralized game theoretic approach for resource allocation of the primary and secondary users in a cognitive radio network is proposed by relaxing these assumptions and defining priority levels. The priorities of the networks are incorporated in the proposed utility and potential functions. It is demonstrated analytically and through simulation studies that a unique NE exists for the combined game with primary users and secondary users (SU). Since the interaction among the networks and their priority levels are incorporated in the functions, the proposed game can be extended to a game among multiple co-existing networks, each with different priority levels.

The contributions of the work include the incorporation of priority levels in utility and potential functions, analytical proofs of existence and uniqueness of NE for the PU and SUs by using these functions with priority levels, and existence of the NE for the combined game. The net result is a game theoretic approach applicable to multiple co-existing networks. Therefore in this paper, after presenting the system representation of a cognitive network in Section 2, the individual games for primary and secondary networks are proposed, and the existence and uniqueness of NE for each of the networks is shown in Section 3. Then by incorporating priority parameters, these two games are combined and the existence of NE for the combined game is proven through exact potential function. Section 4 presents the simulation results, showing the convergence of the games and effect of priority parameter. Finally, Section 5 concludes the paper.

2 System Model

We consider a cognitive radio network consisted of a network of primary users (PUs), and a network of secondary users (SUs). Given a heterogeneous network, it is

assumed that the wireless devices are cognitive and capable of sensing the environment and adapting their parameters. An example of the cognitive network is illustrated in Fig. 1.

In this paper, M PU nodes and N SU nodes with Q orthogonal channels are considered. The transmission power at the i th PU is denoted by p_i^P and the transmission power at the j th SU is given by p_j^S while it is assumed that the SU transmitter/receiver pairs are within the communication range of each other. The communication between the SU or PU pairs can experience interference from transmissions emanating from other PUs or SUs that are using the same channel, and are within the sensing range of the receivers.

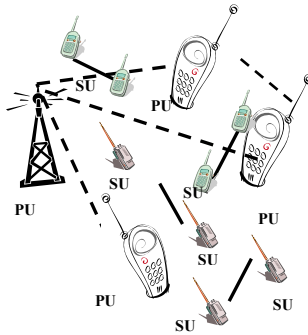


Fig. 1. Cognitive radio network architecture

Given the channel and transmission power as the network resources, the objective is to allocate these resources such that *a)* the interference that SUs cause to the PUs will be minimized, and *b)* SUs be able to communicate to each other. It must be noted that PU nodes do have a very little incentive in reducing their interference on the SU nodes. The objective of the PU network is to satisfy its QoS requirements with the minimum energy consumption. On the other hand, the SU nodes should avoid adding interference to the PU nodes while trying to find a spectrum hole for their own communications.

3 Power Control and Channel Allocation Game

In this section the PU and SU games and their utility functions are introduced. The priorities of the PU and SU networks are incorporated in the utility functions such that the SU nodes receive a larger penalty for causing interference to the PU nodes. On the other hand, the penalty that the PU nodes receive for causing interference to the SU communications is not significant. Subsequently, the existence and uniqueness of NE for the PU and SU games are proven and then the existence of the NE for the combined game is shown by presenting an exact potential function. In the following, the PU and SU games are introduced and existence of NE is demonstrated.

3.1 Primary User Game

Consider a normal¹ form game represented as

$$\Gamma^P = \left\{ M, \{A_i^P\}_{i \in M}, \{u_i^P\}_{i \in M} \right\} \quad (1)$$

where M is the finite set of players, i.e., the PUs. A_i^P is the set of actions available for PU player i , and u_i^P is the utility function associated with player i . If all the available actions for the PU players are collected in action space and denoted by $A^P = \times_{i \in M} A_i^P = A_1^P \times A_2^P \times \dots \times A_M^P$, then $u_i^P : A^P \rightarrow \mathbb{R}$. The decision making by players is accomplished in a distributed manner, but at the same time the decision of each player affects the other players' decisions. The goal of the game is to achieve a set of actions for the players in order to attain an equilibrium point such that no player has any interest in choosing another action unilaterally.

Each PU player action consists of: (a) transmission power, p_i^P , and (b) channel, c_i^P , where $c_i^P \in \{c_1, c_2, \dots, c_Q\}$ represents the channel that PU _{i} is using (superscript P denoting primary users), and Q represents the number of channels. Let the utility function for the PU player i be defined as

$$u_i^P(a_i, a_{-i}) = \alpha_p \log_2(1 + p_i^P h_{ii}^P) - \beta_p \left(\sum_{k \neq i}^M p_k^P h_{ki}^P e(c_k^P, c_i^P) + \sum_{k=1}^N p_k^S h_{ki}^S e(c_k^S, c_i^P) \right) \\ - \left(\eta_p \sum_{k \neq i}^M p_i^P h_{ik}^P e(c_i^P, c_k^P) + \lambda_p \sum_{k=1}^N p_i^P h_{ik}^P e(c_i^P, c_k^S) \right) - \kappa_p p_i^P \left(\sum_{k \neq i}^M p_k^P + \sum_{k=1}^N p_k^S \right) \quad (2)$$

where h_{ij}^P is the channel gain between the transmitter PU node i and receiver j , h_{ij}^S is the channel gain between the transmitter SU node i and receiver node j , and $e(\cdot)$ is a

Boolean function operator as $e(x, y) = \begin{cases} 1 & \text{if } x = y \\ 0 & \text{if } x \neq y \end{cases}$. In the utility function, the first

term is associated with the payoff at the receiver node of PU transmitter i while the second term represents the interference that other nodes cause at PU _{i} . The third term accounts for the interference that PU _{i} causes at the other PU players and also at the SU players whereas the fourth term accounts for the price (penalty) that the user pays for consuming more energy. Note that parameters $\alpha_p, \beta_p, \eta_p, \lambda_p$ and κ_p are the weights that determine the effect of each term in the utility function. In Section 3.4, the constraints on these parameters are discussed to satisfy the uniqueness of Nash Equilibrium and existence of Nash Equilibrium for the potential game.

Remark 1. It must be noted that the summation of the first and the second terms resembles the Shannon capacity [15] with a twist. The interference term has been moved out of the \log term so that the required conditions for the utility functions and potential functions (Sections 3.3, 3.4 and 3.5) will be satisfied.

¹ A normal (strategic) form game is a game with three elements: the set of players, $i \in M$, the strategy space A_i for each player i , and payoff functions that give player i 's utility u_i for each profile of strategies.9.

3.2 Secondary User Game

Consider a normal form game represented as

$$\Gamma^S = \left\{ N, \{A_i^S\}_{i \in N}, \{u_i^S\}_{i \in N} \right\} \quad (3)$$

where N is the finite set of players, i.e., the SUs, A_i^S is the set of actions available for SU player i , and u_i^S is the utility function associated with player i . If all available actions for the SU players are collected in action space and denoted by $A^S = \times_{i \in N} A_i^S$, then $u_i^S : A^S \rightarrow \mathbb{R}$. The decision making by players is distributed similar to the PU and it affects the other players' decisions. The goal of the SU game is to achieve a set of actions for the players yielding an equilibrium point such that no player has any interest in choosing another action unilaterally. The action for each SU player is consisted of: (a) transmission power, p_i^S , and (b) channel, c_i^S , where $c_i^S \in \{c_1, c_2, \dots, c_Q\}$, represents the channel that SU $_i$ is using (superscript S denoting secondary users), and Q represents the number of channels.

Let the utility function for SU player i be defined as

$$u_i^S(a_i, a_{-i}) = \alpha_s \log_2 \left(1 + p_i^S h_{ii}^S \right) - \beta_s \left(\sum_{k=1}^M p_k^P h_{ki}^P e(c_k^P, c_i^S) + \sum_{k \neq i}^N p_k^S h_{ki}^S e(c_k^S, c_i^S) \right) \\ - \left(\eta_s \sum_{k=1}^M p_i^S h_{ik}^S e(c_i^S, c_k^P) + \lambda_s \sum_{k \neq i}^N p_i^S h_{ik}^S e(c_i^S, c_k^S) \right) - \kappa_s p_i^P \left(\sum_{k=1}^M p_k^P + \sum_{k \neq i}^N p_k^S \right) \quad (4)$$

where h_{ij}^P and h_{ij}^S are the channel gains where the transmitter node is PU and SU, respectively, and $e(\cdot)$ is the Boolean function - which all were defined earlier.

In this utility function, the first term is associated with the payoff at the receiver node of SU transmitter i while the second term represents the interference that SU $_i$ causes at the PU players the other SU players. The third term accounts for the interference that SU $_i$ causes at the PU players and also at the other SU players whereas the fourth term accounts for the price that it pays for consuming more energy. Note that parameters $\alpha_s, \beta_s, \eta_s, \lambda_s$ and κ_s are the weights that determine the effect of each term in the utility function. Next it will be shown that NE exists and it is unique.

3.3 Existence of Nash Equilibrium (NE)

In this section the existence of NE for the PU and SU games is demonstrated. Before we proceed, the following is needed.

Definition 1. The normal game $\Gamma = \left\{ N, \{A_i\}_{i \in N}, \{u_i\}_{i \in N} \right\}$ has a NE if for all $i \in N$,

$\{A_i\}_{i \in N}$ is a nonempty convex compact subset of a Euclidean space, and $u_i(a)$ is continuous and concave/quasiconcave on A_i [9].

Next, the following theorem is introduced.

Theorem 1. The PU game $\Gamma^P = \{M, \{A_i^P\}_{i \in M}, \{u_i^P\}_{i \in M}\}$ has a NE.

Proof. In the game $\Gamma^P = \{M, \{A_i^P\}_{i \in M}, \{u_i^P\}_{i \in M}\}$, $\{A_i^P\}_{i \in M}$ is a nonempty compact subset of Euclidean space $\mathbb{R} \times \mathbb{N}$. Also recall (2) where the utility function is continuous in A_i^P . The first derivative of the utility function is given by

$$\frac{\partial u_i^P}{\partial p_i^P} = \frac{\alpha_p}{\ln 2} \cdot \frac{h_{ii}^P}{1 + p_i^P h_{ii}^P} - \left(\eta_p \sum_{k \neq i}^M h_{ik}^P e(c_i^P, c_k^P) + \lambda_p \sum_{k=1}^N h_{ik}^P e(c_i^P, c_k^S) \right) - \kappa_p \left(\sum_{k \neq i}^M p_j^P + \sum_{k=1}^N p_k^S \right) \quad (5)$$

Consequently, the second derivative of the utility function is given by $\partial^2 u_i^P / \partial p_i^P \partial p_i^P = -(\alpha_p / \ln 2) \cdot (h_{ii}^P)^2 \cdot (1 + p_i^P h_{ii}^P)^{-2}$. Hence $\partial^2 u_i^P / \partial p_i^P \partial p_i^P < 0$, and it can be concluded that the utility function, u_i^P , is concave on P_i . Recalling *Definition 1*, we have shown that NE exists for the PU game Γ^P . ■

Theorem 2. The SU game $\Gamma^S = \{N, \{A_i^S\}_{i \in N}, \{u_i^S\}_{i \in N}\}$ has a NE.

Proof. Follow similar steps as the PU game. ■

3.4 Uniqueness of Nash Equilibrium

The best-response correspondence² of the PU and SU games were found in the previous section. However, they must satisfy certain conditions so that a unique NE can be attained for the PU and SU games. First a definition of a unique NE is introduced followed by a theorem to show that indeed the NE for the PU and SU are unique.

Definition 2. A Nash equilibrium is unique if the best-response correspondence, $r(\mathbf{p})$, is a *type II standard function* [13] with following properties: (a) *Positivity*: $r(\mathbf{p}) > 0$, (b) *Type II Monotonicity*: if $\mathbf{p} \leq \mathbf{p}'$, then $r(\mathbf{p}) \geq r(\mathbf{p}')$, and (c) *Type II Scalability*: for all $\omega > 1$, $r(\omega \mathbf{p}) \geq (1/\omega)r(\mathbf{p})$. ■

Theorem 3. NE for the PU game and SU game are unique.

Proof. The uniqueness of the NE for the PU and SU games is demonstrated by showing that the best-response correspondences for them are type II standard functions.

Uniqueness of NE for the PU game. The best-response correspondence of the PU game is found by using $\partial u_i^P / \partial p_i^P = 0$. Recalling (5), the best-response correspondence is given by

$$r_i^P(\mathbf{p}) = p_i^P = \frac{\alpha_p}{\ln 2} \cdot \left(\eta_p \sum_{k \neq i}^M h_{ik}^P + \lambda_p \sum_{k=1}^N h_{ik}^P + \kappa_p \left(\sum_{k \neq i}^M p_j^P + \sum_{k=1}^N p_k^S \right) \right)^{-1} - \frac{1}{h_{ii}^P}. \quad (6)$$

² Best-response correspondence (function) for player i is a function that yields a set of actions for player i , $r_i(a_{-i})$, when the other players' actions are given by a_{-i} , i.e. $r_i(a_{-i}) = \{a_i \text{ in } A_i : u_i(a_i, a_{-i}) \geq u_i(a'_i, a_{-i}) \text{ for all } a'_i \text{ in } A_i\}$. Any action in that set is at least as good for player i as every other action of player i , when the other players' actions are given by a_{-i} . 9.

However, the coefficients must satisfy certain conditions so that a positive real value can be achieved for p_i^P that maximizes the utility function. This p_i^P is the unique NE.

It can be proved that the three properties of type II standard function are satisfied for the best-response correspondence function (the proofs omitted due to page limit). Therefore, the best-response correspondence is a type II standard function. Recalling *Definition 2*, it can be concluded that the NE obtained by the best-response correspondence is unique [13].

Uniqueness of NE for the SU game. The uniqueness of NE for the SU game $\Gamma^S = \{N, \{A_i^S\}_{i \in N}, \{u_i^S\}_{i \in N}\}$ can be shown similar to the PU game. ■

In this section we showed that the PU and SU games have unique NE. In the next section, the combined game is considered and existence of the NE is demonstrated.

3.5 Existence of Nash Equilibrium - Exact Potential Function

In this section, the combined PU and SU game is considered, and the existence of NE for the combined game is shown. This is done by proving that the combined game is a potential game. A potential game [11] is a normal game such that any change in the utility function of a player caused by a unilateral change in the player’s strategy, is reflected in a potential function which represents the aggregated payoff of all the players. In other words, any single player’s interest in increasing its own payoff is aligned with the group’s interest. Furthermore, Hofbauer and Sandholm in [12] presented that “in potential games, the aggregate payoff function (i.e. potential function) serves as a Lyapunov function for a broad class of evolutionary dynamics.”

Definition 3. A game $\Gamma = \{N, \{A_i\}_{i \in N}, \{u_i\}_{i \in N}\}$ is a *potential game* [11] if there is a potential function $V : A \rightarrow \mathbb{R}$ such that one of the following conditions holds. A game satisfying the first condition is an *exact potential game*, and a game satisfying the second condition is an *ordinal potential game*.

- i) $V(a_i, a_{-i}) - V(a'_i, a_{-i}) = u(a_i, a_{-i}) - u(a'_i, a_{-i})$, for $\forall i \in N, a \in A$, and $a'_i \in A_i$.
- ii) $u(a_i, a_{-i}) > u(a'_i, a_{-i}) \Leftrightarrow V(a_i, a_{-i}) > V(a'_i, a_{-i})$, for $\forall i \in N, a \in A$, and $a'_i \in A_i$.

Theorem 4. The combined game of PU and SU converges to NE for all the players.

Proof. If the game can be shown to be an exact potential game (EPG) with a potential function, then it has been proven that the game strategy towards the *best response* converges to NE [11].

Consider the following potential function candidate

$$\begin{aligned}
 V(a) = & \alpha_P \sum_{i=1}^M \log_2(1 + p_i^P h_{ii}^P) + \alpha_S \sum_{i=1}^N \log_2(1 + p_i^S h_{ii}^S) - a\beta_P \sum_{j=1}^M \left(\sum_{k \neq j} p_k^P h_{kj}^P e(c_k^P, c_j^P) + \sum_{k=1}^N p_k^S h_{kj}^S e(c_k^S, c_j^P) \right) \\
 & - b\beta_S \sum_{j=1}^N \left(\sum_{k=1}^M p_k^P h_{kj}^P e(c_k^P, c_j^S) + \sum_{k \neq j} p_k^S h_{kj}^S e(c_k^S, c_j^S) \right) - c \sum_{j=1}^M \left(\eta_P \sum_{k \neq j} p_j^P h_{jk}^P e(c_j^P, c_k^P) + \lambda_P \sum_{k=1}^N p_j^P h_{jk}^P e(c_j^P, c_k^S) \right) \\
 & - d \sum_{j=1}^N \left(\eta_S \sum_{k=1}^M p_j^S h_{jk}^S e(c_j^S, c_k^P) + \lambda_S \sum_{k \neq j} p_j^S h_{jk}^S e(c_j^S, c_k^S) \right)
 \end{aligned}$$

$$-e\kappa_p \sum_{j=1}^M p_j^p \left(\sum_{k \neq j}^M p_k^p + \sum_{k=1}^N p_k^s \right) - f\kappa_p \sum_{j=1}^N p_j^s \left(\sum_{k=1}^M p_k^p + \sum_{k \neq j}^N p_k^s \right). \quad (7)$$

It can be shown that function $V(a)$ is an exact potential function for PU and SU games if the following conditions are satisfied for $a, b, c, d > 0$:

$a = b, c = d = 1 - a, e = f = 1/2, \beta_p = \eta_p = \eta_s, \beta_s = \lambda_s = \lambda_p$, and $\kappa_p = \kappa_s$, and the PU and SU utility functions are defined as

$$\begin{aligned} u_i^p(a_i, a_{-i}) &= \alpha_p \log_2 \left(1 + p_i^p h_{ii}^p \right) - \kappa_p p_i^p \left(\sum_{j \neq i}^M p_j^p + \sum_{k=1}^N p_k^s \right) \\ &- \beta_p \left(\sum_{k \neq i}^M p_k^p h_{ki}^p e(c_k^p, c_i^p) + \sum_{k=1}^N p_k^s h_{ki}^s e(c_k^s, c_i^p) \right) - \left(\beta_p \sum_{k \neq i}^M p_i^p h_{ik}^p e(c_i^p, c_k^p) + \beta_s \sum_{k=1}^N p_i^p h_{ik}^p e(c_i^p, c_k^s) \right) \\ u_i^s(a_i, a_{-i}) &= \alpha_s \log_2 \left(1 + p_i^s h_{ii}^s \right) - \kappa_p p_i^p \left(\sum_{k \neq i}^M p_k^p + \sum_{k=1}^N p_k^s \right) \\ &- \beta_s \left(\sum_{k=1}^M p_k^p h_{ki}^p e(c_k^p, c_i^s) + \sum_{k \neq i}^N p_k^s h_{ki}^s e(c_k^s, c_i^s) \right) - \left(\beta_p \sum_{k=1}^M p_i^s h_{ik}^s e(c_i^s, c_k^p) + \beta_s \sum_{k \neq i}^N p_i^s h_{ik}^s e(c_i^s, c_k^s) \right) \end{aligned}$$

The potential function, $V(a)$, can be written with respect to a PU as

$$V(a_i, a_{-i}) = u_i^p(a_i, a_{-i}) + F^p(a_{-i}), \quad (8)$$

where $F^p(a_{-i})$ is a function independent of the strategy of PU player i , a_i .

Hence, if PU player i changes its strategy to a'_i , the potential function can be written as

$$V(a'_i, a_{-i}) = u_i^p(a'_i, a_{-i}) + F^p(a_{-i}). \quad (9)$$

Consequently,

$$V(a_i, a_{-i}) - V(a'_i, a_{-i}) = u_i^p(a_i, a_{-i}) - u_i^p(a'_i, a_{-i}). \quad (10)$$

Therefore, the potential function is an exact potential function for PUs.

Similarly, the potential function, $V(a_i, a_{-i})$, can be written with respect to a SU as $V(a_i, a_{-i}) = u_i^s(a_i, a_{-i}) + F^s(a_{-i})$, where $F^s(a_{-i})$ is a function independent of the strategy of SU player i , a_i . Consequently, it can be shown that $V(a_i, a_{-i}) - V(a'_i, a_{-i}) = u_i^s(a_i, a_{-i}) - u_i^s(a'_i, a_{-i})$. Therefore, the potential function is exact for SUs. ■

Remark 2. The priorities of the networks are incorporated in the utility functions as the β_p/β_s ratio.

Remark 3. The coefficients β_p and β_s can be chosen such that term associated with the interference caused by the PUs at SUs is less significant than the term associated with the interference caused by the SUs at PUs. It can be achieved by selecting $\beta_p \gg \beta_s$.

Definition 4. P is a potential for game $\Gamma = \{N, \{A_i\}_{i \in N}, \{u_i\}_{i \in N}\}$ if and only if P is differentiable, and $\partial u_i / \partial a_i = \partial P / \partial a_i$ for all $i \in N$ [11].

Theorem 5. V is a potential for the PU game, Γ^P , and SU game Γ^S .

Proof. Recall (8), where the potential function V can be written as $V(a_i, a_{-i}) = u_i^P(a_i, a_{-i}) + F^P(a_{-i})$. The first derivative of the potential function will be $\frac{\partial V}{\partial p_i^P} = \frac{\partial u_i^P}{\partial p_i^P} + \frac{\partial F^P}{\partial p_i^P} = \frac{\partial u_i^P}{\partial p_i^P} + 0 = \frac{\partial u_i^P}{\partial p_i^P}$. Therefore, V is a potential for the PU game.

It can be similarly shown that V is a potential for the SU game. ■

Definition 5. A game $\Gamma = \{N, \{A_i\}_{i \in N}, \{u_i\}_{i \in N}\}$ is an exact potential game if and only if $\partial^2 u_i / \partial a_i \partial a_j = \partial^2 u_j / \partial a_i \partial a_j$ for all $i, j \in N$ [11].

Theorem 6. The PU and SU games are exact potential games.

Proof. For both PU and SU games it can be shown that $\frac{\partial^2 u_j^P}{\partial p_i \partial p_j} = \frac{\partial^2 u_i^P}{\partial p_i \partial p_j} = \frac{\partial^2 u_j^S}{\partial p_i \partial p_j} = \frac{\partial^2 u_i^S}{\partial p_i \partial p_j} = -\kappa$. Therefore, both PU and SU games are exact potential games [11]. ■

4 Simulation Results

In this section, simulation results for a network of 5 PU nodes and 30 SU nodes are presented. Simulations were performed in Matlab®. We assume that the transmission power at the nodes can vary in the range $[0, P_{max}]$ mW. The initial transmission power at the PU and SU nodes were selected as random numbers uniformly distributed between 0 and P_{max} . The weighting factors were considered as following: $\alpha_p = 400$, $\beta_p = 10$, $\alpha_s = 370$, $\beta_s = 1$, $P_{max} = 1$ mW (or 0dBm), $a = 2$, and $\kappa = 1 / ((N + M) \cdot P_{max})$. Fig. 2 shows the transmission power and utility at the PU nodes and five SU nodes as the game is played, and the players update their transmission powers using (6). Though there are 30 SU nodes, we are presenting the results for 5 nodes for the sake of brevity. It can be noticed how each PU and SU player enhances its own utility function by selecting the proper transmission power. Note that since we have chosen $\beta_p = 10\beta_s$, the PU nodes care less about the interference that they cause on the SU nodes. Also it must be noted that since $\alpha_s < \alpha_p$, the utility of SUs eventually becomes less than the utility of PU although they have been assigned the same power.

Fig. 3 shows the potential function of the combined PU and SU game. It can be noticed how the potential function converges to the Nash equilibrium as the PU and SU games (Fig. 2) converge to their Nash equilibria.

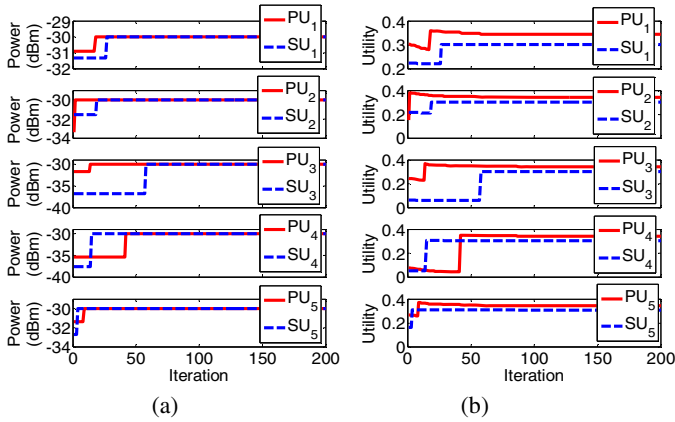


Fig. 2. (a)Transmission power and (b) utility at PU nodes and five SU nodes

In the following we are going to examine the effect of the β_p/β_s ratio on the results. Fig. 4 shows the final utility of the PU and SU nodes when $\beta_p = \beta_s$, $\beta_p = 10\beta_s$ and $\beta_p = 100\beta_s$. The thick markers show the center of the cluster for each case, and the thin markers represent the utility value for each individual node. The PU nodes in each case achieve very similar values of utility, so the individual points cannot be detected in the figure. Note that $\beta_p = \beta_s$ means that the primary and secondary networks have equal rights in using the spectrum and resources. In other words, the combined game will treat the nodes equally, and the interference that they cause on each other will be penalized equally. It can be noticed that when $\beta_p = \beta_s$, the utility of the PU and SU nodes have achieved the mean steady-state values (center of the cluster) of the same range. As β_p/β_s ratio increases, the gap between the PU and SU utility clusters increases. At $\beta_p = 100\beta_s$, the gap between the SU and PU utility clusters is even more prominent where the PU network has a much greater priority than SU network.

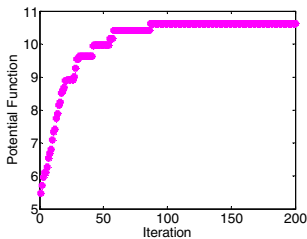


Fig. 3. Potential function, V , of the combined PU and SU game

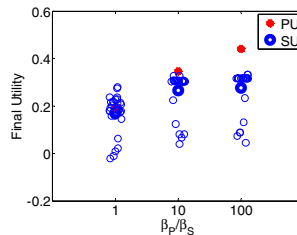


Fig. 4. Final utility at PU and SU nodes for various values of β_p/β_s

5 Conclusions

In this paper we presented a game theoretic approach for resource allocation in primary and secondary users in cognitive networks. The solution takes into account the priority of the networks in their access to the resources. The existence and uniqueness of the Nash equilibrium for the PU and SU games were proven and then a potential function was presented for the combined PU and SU game. Consequently the existence of the Nash equilibrium for the potential game was shown.

The proposed approach provides the flexibility of adjusting to various priorities among the several co-existing networks, while guaranteeing the existence and uniqueness of Nash equilibrium and consequently achieving it. The simulation results show the convergence of the transmission power and utility for the PU and SU users, and also the convergence of the potential function. They also provide results for various cases of network priorities.

References

1. Mitola III, J.: Cognitive Radio: an Integrated Agent Architecture for Software Defined Radio. PhD thesis, KTH Royal Institute of Technology, Stockholm, Sweden (2000)
2. Wang, B., Wu, Y., Liu, K.J.R.: Game Theory for Cognitive Radio Networks: an Overview. *Computer Networks* (2010), doi:10.1016/j.comnet.04.004
3. Giupponi, L., Ibars, C.: Distributed Cooperation among Cognitive Radios with Complete and Incomplete Information. *EURASIP Journal on Advances in Signal Processing*, Article ID 905185, 13 pages (2009)
4. Menon, R., MacKenzie, A.B., Buehrer, R.M., Reed, J.H.: WSN15-4: a Game-Theoretic Framework for Interference Avoidance in Ad Hoc Networks. In: *IEEE Global Telecomm. Conference 2006*, pp. 1–6. IEEE Press, New York (2006)
5. Zhang, Z., Shi, J., Chen, H.-H., Guizani, M., Qiu, P.: A Cooperation Strategy Based on Nash Bargaining Solution in Cooperative Relay Networks. *IEEE Trans. on Vehicular Technology* 57(4), 2570–2577 (2008)
6. Tianming, L., Jayaweera, S.K.: A Novel Primary-Secondary User Power Control Game for Cognitive Radios with Linear Receivers. In: *IEEE Military Comm. Conference*, pp. 1–7. IEEE Press, New York (2008)
7. Ghosh, C., Agrawal, D.P., Rao, M.B.: Channel Capacity Maximization in Cooperative Cognitive Radio Networks Using Game Theory. *SIGMOBILE Mob. Comput. Commun. Rev.* 13(2), 2–13 (2009)
8. Saraydar, C.U., Mandayam, N.B., Goodman, D.J.: Efficient Power Control via Pricing in Wireless Data Networks. *IEEE Trans. on Comm.* 50(2), 291–303 (2002)
9. Osborne, M.J., Rubinstein, A.: *A Course in Game Theory*. The MIT Press, Cambridge (1994)
10. Meshkati, F., Poor, H.V., Schwartz, S.C., Balan, R.V.: Energy-Efficient Resource Allocation in Wireless Networks with Quality-of-Service Constraints. *IEEE Trans. on Comm.* 57(11), 3406–3414 (2009)
11. Monderer, D., Shapley, L.S.: Potential Games. *Games and Economic Behavior* (14), 124–143 (1996)
12. Hofbauerand, J., Sandholm, W.: On the Global Convergence of Stochastic Fictitious Play. *Econometrica* 70, 2265–2294 (2002)
13. Sung, C.W., Leung, K.-K.: A Generalized Framework for Distributed Power Control in Wireless Networks. *IEEE Trans. on Info. Theory* 51(7), 2625–2635 (2005)
14. Fudenberg, D., Tirole, J.: *Game Theory*. MIT Press, Cambridge (1991)
15. Stallings, W.: *Wireless Communications & Networks*. Prentice Hall, USA (2005)

A Method for Identifying and Analyzing Authentication Mechanisms and a Case Study

Sérgio Ribeiro¹ and Danilo Suiama²

¹ Fundação CPqD, Campinas, Brazil
sribeiro@cpqd.com.br

² Elektro, Campinas, Brazil
Danilo.Suiama@elektro.com.br

Abstract. This paper presents a method able to identify and analyze authentication mechanisms for a given type of application. Currently, there are several **applications** that use and require a mechanism to provide the user with physical or logical access. This article proposes a solution developed by Fundação CPqD that identifies which mechanism is more suitable and recommended for a given type of application, considering the restrictions and challenges for every type of mechanism. Finally, a case study will be exposed providing an example of authentication mechanism analysis and identification system.

Keywords: Methodology, Authentication Mechanisms.

1 Introduction

The high level of connectivity, the constant need to access information, the mobility and the ever-present changes in user behavior imposed by technological innovations come together as a security challenge to be overcome: to identify the most appropriate authentication mechanism for a specific application.

The BIOMODAL Project is funded by the Brazilian Innovation Agency (FINEP), subordinated to the Ministry of Science and Technology, and the Telecommunications Technology Development Fund (FUNTTEL), subordinated to the Ministry of Communications. Its main objective is to implement iconographic and multimodal biometric authentication on mobile devices. Several applications require authentication for the logical and physical access control and this has brought up the need to develop a mechanism for identifying and analyzing the most appropriate authentication mechanism for a given context. Therefore, this paper will present the method that was developed.

In general, the method will analyze all aspects such as the challenges offered by the authentication mechanisms and will show the most appropriate solution for a given application context. For instance, one of the problems is the amount of passwords required to access the countless applications, which affect other services. This happens when using weak passwords easy to be memorized for multiple accesses, or when using few passwords to access several types of services. According to Carrie-Ann (2009), this is a real problem, as around one third of the users use a single password for different services, whereas only 19% of users do not use the same passwords.

The role of the method and a case study involving the authentication mechanism analysis and identification system will examine aspects such as simplicity, fraud (stealing), loss, forgetting the password, and other authentication problems.

2 MIAMA Structure

MIAMA – Method for Identifying and Analyzing Authentication Mechanisms¹ – is composed of nine phases. The result will map and establish the authentication mechanisms for the required context.

The following figure 1 shows the six macro processes that compose the MIAMA structure: Identification, Definition, Assignment, Analysis, Mapping and Priority. The figure also shows the main objectives of each phase and its relationship to other phases.

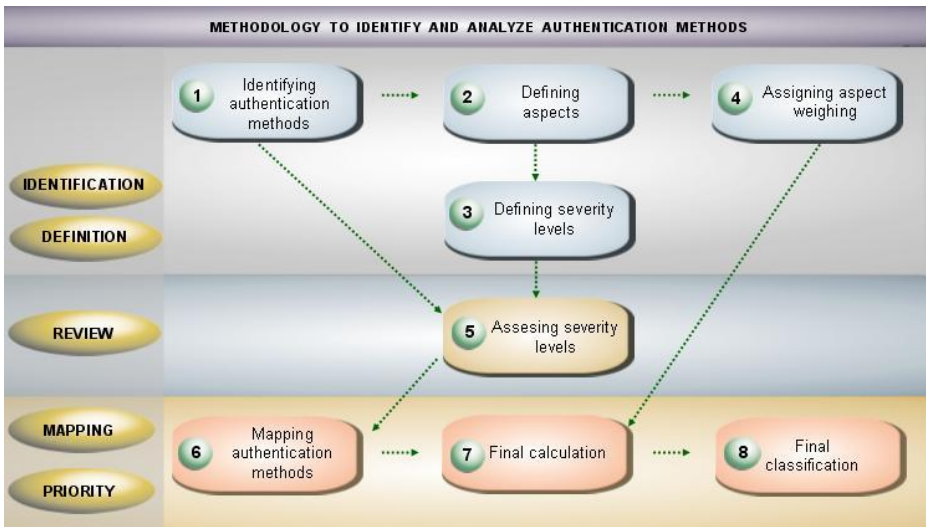


Fig. 1. Phase Chart

2.1 Context Identification

To establish the most appropriate authentication, the first phase of the method is to identify the application context. Currently, there are several applications that require authentication for the logical and physical control. This phase identifies the scenario where the authentication mechanism will be applied.

These are some sectors where the authentication applications can be applied: Finance, corporate, business, etc. The main objective is to understand the characteristics of each sector and mainly the reasons to select a given authentication mechanism.

The result obtained from this phase will be the basis for the next phases. At this stage, the aspects will be analyzed and the weight to be assigned to these aspects according to the identified context will be established.

¹ The acronyms shown are derived from the name in Portuguese.

2.2 Identifying the Authentication Mechanisms

After identifying the application context, the selection of authentication mechanisms to be analyzed will happen in the second phase of the method.

This phase will identify the authentication mechanisms to be analyzed. The mechanisms can be stated as:

- a. Biometric-based authentication (digital, retina, iris, voice, face, DNA, etc.);
- b. Token-based authentication (token, identification card, passport, etc.);
- c. Knowledge-based authentication (passwords, security phrases, PIN, etc.).

2.3 Defining the Aspects

Once the application context and the authentication mechanisms are established, the next stage will be to define the important aspects to compare and analyze the authentication mechanisms.

In this method elaboration exercise, some strategic aspects were established for its validation, however, not restricted to the listed set of mechanisms. Therefore, MIAMA can establish the following set of aspects: Universality, Acceptability, Convenience, Simplicity, Intrusive, Privacy, Reliability, Security level, Reissuing/Revoking, Duration, Processing, Precision, Interference, Collectibility, Research and Development, Products, Services and Cost.

2.4 Defining the Severity Levels

The objective of this phase is to establish the qualitative severity levels associated with each aspect listed in Phase 3 (Aspect Definition) and the corresponding criteria in order to evaluate the authentication mechanisms established in Phase 2 (Authentication Mechanism Identification).

The severity levels establish the influence that each authentication mechanism has on a given aspect. For this reason, the criteria used to establish the severity levels should consider not only the global aspects listed in Phase 3 (Aspect Definition), but also a specific analysis of every aspect, as each level varies according to the aspect to be analyzed. In this way, every aspect must be analyzed and its severity level defined according to what was established during the method phase.

For example, three severity levels associated with each aspect were set as “High”, “Average” and “Low”.

2.5 Assigning the Aspect Weighing

In order to differentiate the importance of each aspect, the weight of each aspect is established in this phase. It is important to state that Phase 1 has a direct impact on this stage. The definition of the application context is an important point to establish the most relevant aspects.

At first, every weight must be set and validated by a group of specialists, in order to suit the scenario to be investigated. For this reason, this phase will require most of the process time.

2.6 Assessing the Severity Levels

In general, each selected aspect is connected to a score. The criteria can be represented by a binary variable (“Yes” or “No” type) or qualitative variable (“High”, “Average” and “Low”). For every aspect, when the criteria is a binary variable, the maximum preestablished score will be assigned if the variable is “Yes”, and the minimum score if the variable is “No”.

An example using three authentication mechanisms (MA1, MA2 and MA3) in which the severity levels are assigned for each aspect is listed next.

Table 1. Severity Level Assignment

ASPECTS	MA1	MA2	MA3
Aspect 1	No	No	Yes
Aspect 2	Low	Average	High
Aspect 3	Low	Low	High
Aspect 4	Low	Average	High
...	High	High	Low
Aspect <i>n</i>	Low	Average	High

2.7 Mapping the Severity Levels

By assigning the severity levels, the next objective is to quantify the severity degree. For instance, a scale from 1 to 3 can be used, where the maximum score will be “High” or “Yes”.

Table 2. Severity Level Mapping

ASPECTS	MA1	MA2	MA3
Aspect 1	1	1	3
Aspect 2	1	2	3
Aspect 3	1	1	3
Aspect 4	1	2	3
...	3	3	1
Aspect <i>n</i>	1	2	3

2.8 Final Calculation

In order to consolidate the assigned weight and scores, the equation (1) will be used to calculate the final score of each authentication mechanism. In general, this equation contains a function that adds the scores from the aspects and their weight.

$$PMa_n = \sum_{i=1}^K \alpha_i N_i \quad (1)$$

where:

PMa_n = Final result (points) from the evaluated authentication mechanism n ;

K = Number of criteria;

N_i = Score assigned to an aspect i ;

α_i = Weight of aspect i .

2.9 Final Classification

Finally, the last phase of the method will be to understand the total score and the classification so that the authentication mechanisms are applied to a given application context.

3 MIAMA Application

To automate the analysis process of authentication mechanisms and classify them, a system with Web access was selected and deployed to record the data and analyze the results, which will facilitate the entire process.

The tool basically allows two ways of interaction, one manual that will be explained in the following sections, and another automated. In the latter, the user will import the data from a spreadsheet file. During the import process, all data will be validated and listed according to the method. After loading the data, the tool will calculate the score of each authentication mechanism included in the file.

The user will be able to define a new list of criteria scores for every aspect, in addition to the weight, which will have an impact on the selected context. This will facilitate simulations and provide a better understanding of the subject. In case a score or weight is changed, the tool will perform a new calculation of the authentication mechanisms, providing more flexibility in both the analysis and value of the criteria.

In addition to that, another facility provided by the tool is the use of filters. Through the filters, the user can perform the following queries: for instance to identify the scores by criteria, or list the scores by aspect.

3.1 Authentication Mechanism Analysis and Identification System

The following sections describe the way SIAMA – Authentication mechanism Analysis and Identification System² – registers and analyzes user decision regarding each included aspect.

² The acronyms shown are derived from the names in Portuguese.

As explained previously, the procedure can be performed manually or automatically. This document presents the main stages of the Manual mode. After performing the process manually, the tool generates a normalized table. The fields from the table can be edited in order to reload and analyze the authentication mechanisms.

3.1.1 Case Study

The case study stated in this paper illustrates the most feasible authentication mechanism for a school canteen. The aim of the canteen is to use a mechanism to help the staff sell food according to the profile of the students. The objective of this restriction measure to sell food and drinks in the school environment is the adaptation to the state law and creation of better eating habits since childhood.

The system process was divided in two parts, as established in the following example of method application simulation: registration and assignment.

The registration includes four phases from the method. The user must enter the data for analysis. For the assignment, the importance/significance of each aspect will be analyzed according to the application context.

3.1.2 Registration

The purpose of the first phase is to include initial data for analysis. The scenario to be applied, the authentication mechanisms, and the main aspects and criteria to be analyzed will be registered in the system.

The following figure shows the first user interface. Here, the application context or probable cases in which the authentication mechanisms are inserted.

The screenshot displays the 'Authentication Method Analysis and Identification System' interface. At the top left is the CPD logo. The main title is 'Authentication Method Analysis and Identification System'. Below this is a section titled 'Application Context Registration'. This section contains two main input areas: 'Context' and 'Description'. The 'Context' field is a dropdown menu currently showing 'School Canteen'. To its right are 'Add' and 'Delete' buttons. The 'Description' field is a text area containing the text: 'The school canteen will use the authentication method. This will help its staff to sell food matching every student's profile.' To its right is a 'Back' button.

Fig. 2. Context registration

After performing the registration of the application context, the authentication mechanisms to be analyzed and compared for the deployment in the school canteen will be inserted.

Figure 3 shows the registration stage. It is possible to add a short little note for each authentication mechanism in the Description field.

The screenshot shows the 'Authentication Method Registration' window. It features a logo on the top left and the title 'Authentication Method Analysis and Identification System'. The main content area is titled 'Authentication Method Registration' and contains two sections: 'Authentication Methods' and 'Description'. The 'Authentication Methods' section has a list box with 'Biometry (general)', 'Password', and 'Token' selected. To its right are 'Add' and 'Delete' buttons. The 'Description' section has a text area containing 'Biometry to be used was not defined' and a 'Back' button to its right.

Fig. 3. Authentication mechanism Registration

Once the authentication methods are defined, the next step is to add the aspects that will be compared and analyzed in order to identify the most appropriate for each context.

To perform an accurate analysis it is important to define these aspects, therefore, it is required to register strategic points that can limit the collection of authentication mechanisms and validate the method. Some aspects and their brief description are illustrated next.

The screenshot shows the 'Aspect Registration' window. It features a logo on the top left and the title 'Authentication Method Analysis and Identification System'. The main content area is titled 'Aspect Registration' and contains two sections: 'Aspect' and 'Description'. The 'Aspect' section has a list box with 'Universality', 'Acceptance', 'Suitability', and 'Ease of Use' selected. To its right are 'Add' and 'Delete' buttons. The 'Description' section has a text area containing 'It is understood that the universality aspect is the characteristic of all users during authentication.' and a 'Back' button to its right.

Fig. 4. Aspect registration

Another important aspect is to establish the severity levels. The result obtained from this stage will be used as a guide for the end result. Then, the qualification criteria must be established in order to avoid possible ambiguity regarding decisions. Usually, this task is supported by experts in the area, in order to minimize possible inappropriate decisions.

CPD Authentication Method Analysis and Identification System

Severity Levels Registration

Aspect

- Universality
- Acceptance
- Suitability
- Ease of Use

High You do not need to load devices for each application

Average Load device

Low Passwords hard to remember and/or loading devices for each application

Back

Fig. 5. Severity Level Registration

3.1.3 Assignment

After performing the data registration, the next stage is to assign the weight and severity level to each aspect through application context. According to the application context, some requirements are considered more important, such as:

- **Acceptability:** The objective of the acceptability aspect is to show the user satisfaction regarding the selected authentication mechanisms. From the user's point of view, this is an important aspect to analyze the acceptability of the mechanism.
- **Simplicity:** The objective of this process is to identify the intricate aspects of using the authentication mechanism. This analysis will validate the main obstacles during user authentication.
- **Convenience:** The objective of this process is to analyze the user effort related to the authentication elements. In other words, an evaluation will be performed whether the user needs to memorize, load or make any effort during authentication.

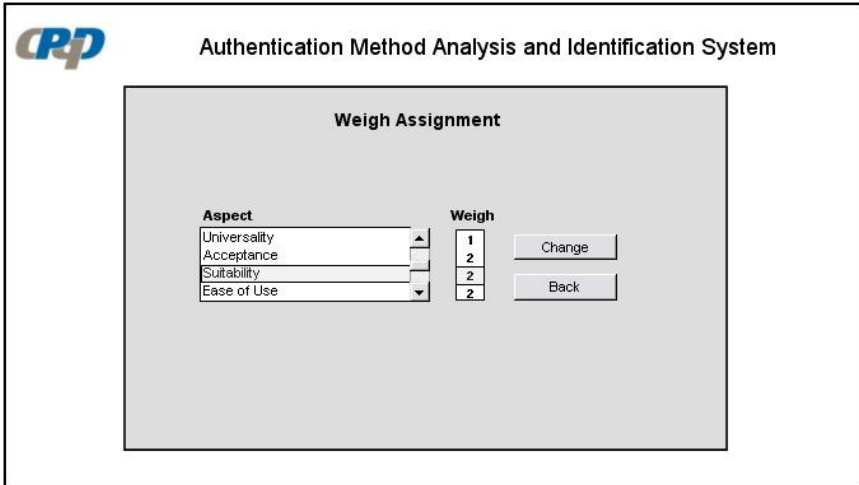


Fig. 6. Weight assignment

When the aspect significance is established, the last stage before the final calculation is to assign the severity levels of each aspect associated with the authentication mechanism. Then, all the aspects of the authentication mechanisms being analyzed must be rated according to the severity level.

The severity level regarding the convenience aspect for the selected authentication mechanism is illustrated next. It is important to state that an average severity level is assigned to the school canteen.

A conclusion was reached that, for the school canteen, the use of tokens is considered to be the average level. This was due to the fact that the student can forget, break or even lend the token to other peers, which will influence in the obtained report results.

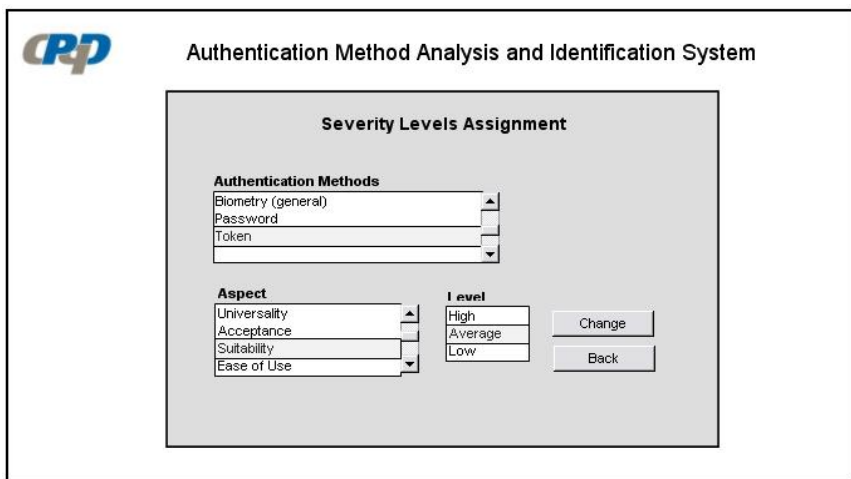


Fig. 7. Severity Level Assignment

Finally, the last stage includes the final classification of the authentication mechanisms. An example of the obtained result for each mechanism and its classification by the system

Note that the most feasible result regarding the school canteen, with restriction measures to sell food and drinks in the school environment, is the use of biometry. This process was selected because it seems to be the most appropriate among the other analyzed authentication mechanisms.

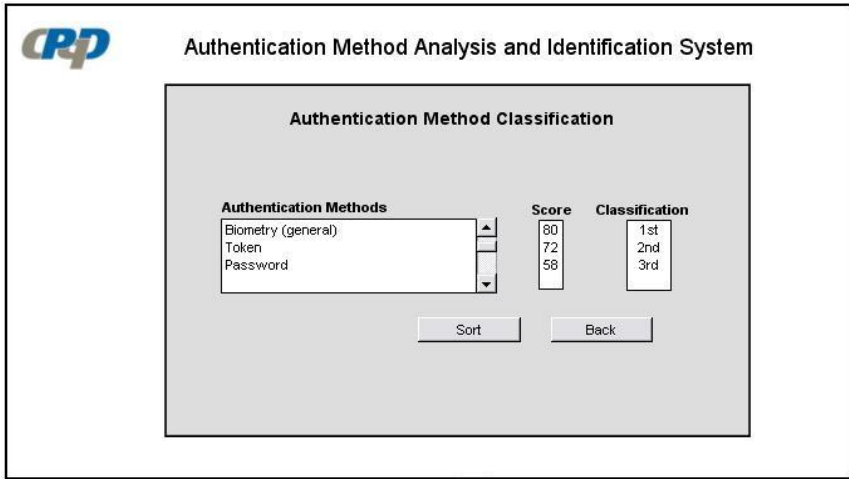


Fig. 8. Authentication mechanism Classification

4 Conclusion

This paper reveals the method to identify and analyze the authentication mechanisms when applied in different contexts, listing the practical and theoretical fields, and the application through the developed system model.

The method is composed by nine phases, in which the focus was given to the analysis of authentication mechanisms according to the required application context. By using this method, the user will be able to understand the authentication mechanisms providing details on the restrictions and challenges.

Conceptually, MIAMA can be applied to several authentication mechanisms. The final result is considered after analyzing all aspects, not restricted to specific selection criteria. Easy to understand and the reason why a specific mechanism can be selected, this method identifies first the application context where the result from the context will guide through the most relevant aspects.

The end result allows the elaboration of strategic plans in order to deploy an authentication mechanism for a specific type of scope. This article also included a case study analyzing a school canteen where several requirements had to be considered. For this scenario, the tool selected Biometry as the most feasible mechanism when compared to tokens and passwords.

References

1. Skinner, C.-A.: PC World: One-Third Use a Single Password for Everything (March 11, 2009), http://www.pcworld.com/article/161078/one_third_use_same_password.html
2. Suiama, D.Y., Ribeiro, S.L.: Metodologia de Identificação e Análise dos Métodos de Autenticação – PD.30.13.22A.0001A-RT36-AA, CPqD, Campinas, Brazil (2010)
3. Tao, H.: Pass-Go, a New Graphical Password Scheme. Thesis submitted to the Faculty of Graduate and Postdoctoral Studie In partial fulfillment of the requirements For the Master of Applied Science degree in Electrical and Computer Engineering. University of Ottawa, Canada (June 2006)
4. Ishikawa, K., (Lu, D.J. trans.): What is Total Quality Control? Prentice-Hall Inc., Englewood Cliffs (1985)
5. Qibin, S., Zhi, L., Xudong, J., Kot, A.: An Interactive and Secure User Authentication Scheme for Mobile Devices. In: IEEE International Symposium on Circuits and Systems (ISCAS), pp. 2973–2976 (May 2008)
6. Saaty, T.L.: The Analytic Hierarchy Process. McGraw Hill, New York (1980)

An AI Based Framework Proposal for Real Time Robot Design Using Framsticks

S. Raja Mohamed¹ and P. Raviraj²

¹ Associate Professor / Dept of CSE
Sri Vishnu Engg College for Women, WG District, AP, India
rajamohameds@svecw.edu.in

² Professor / Dept of CSE
SKP Engg College, Tiruvannamalai, TN, India
raviraj_it@yahoo.co.in

Abstract. Real time modeling of Robot is increasing day by day especially in automated industries. In this paper we propose a framework to design real time robots for industries using AI techniques like Co-Evolution, Virtual Ecology, Life time learning. Beyond that monitoring complex and non complex behaviors in different environments and obtaining the parameters that influence hardware design. Here we have created a virtual khepera robot and simulated in Framsticks and designed hardware components based on the outcome of simulation parameters. Even control programs were also generated using the same data.

Keywords: Artificial Intelligence, Ecology, Khepera, Framsticks.

1 Introduction of Virtual World

The realistic physical modeling of characters in games and virtual worlds is becoming a viable alternative to more traditional animation techniques. Physical modeling of characters and environments can enhance realism and allow users to interact with the virtual world much more freely. By modeling the forces and torques acting on bodies within a virtual environment, detecting and responding to collisions between bodies, respecting constraints arising from, for example, joints in articulated multi-body objects, etc., the system behaves in a believable manner in all situations, and the user can therefore be allowed to interact with it much more freely. The growing popularity of this approach is demonstrated by the appearance over the last couple of years of a number of off-the-shelf physics engines aimed at games programmers [1]. While these engines handle the modeling of inanimate bodies, programmers are still left with the task of writing controllers for motile objects within the environment (e.g. cars, human characters, monsters, etc.). Writing a controller for a physically modeled character is a question of calculating the appropriate forces to apply to each body part at each point of time to achieve the desired movement (e.g. a realistic walking gait for a human character). Artificial life techniques can be useful in automating this task [2] [3]. For example, artificial evolution can generate suitable controllers for simple behaviors, given only a high level description of that behavior in terms of a fitness function. In this paper, the state of the art in evolving controllers, and also in evolving the

characters' body shapes, is described. It is then suggested that current approaches will not be able to scale up to more complicated behaviors.

2 Data Probing

Simulation is a technique for exploring interesting regions of this immense landscape of robot design. It is a platform for generating suitable and interesting forms and behaviors, not limited by the preconceptions of a human designer's imagination. It can only get off the ground, if the initial randomly generated creature with a non-zero score on the fitness function. For example, consider an attempt to evolve a behavior whereby a creature needed to process complex visual data about the movement of another creature standing in front of it, and use this to decide whether that creature is a friend or foe. This enterprise would clearly have little chance of success if the evolutionary process was starting from scratch.

3 Proposed AI Techniques

Researchers in the field of evolutionary robotics have considered various methods of overcoming the difficulties like selecting a non-zero fitness function, mutation etc. These generally involve either the incremental acquisition of increasingly complicated tasks [3], or the decomposition of the task into a sequence of easier tasks together with some way of combining them. A problem with many of these approaches is that the decomposition of the task into easier and/or incremental steps is something of an art in itself and there are no general guidelines to suggest the most appropriate way to do this sensible task. Decomposition from the designer's point of view may not be the best route by which one can evolve a complex behavior. Despite the problems described above with evolving single creature to perform complex tasks, there are a number of alternative approaches that have shown some signs of success. Five such popular methods are given below:

3.1 Co-evolution

Co-evolution is a very promising technique for developing complex behaviors, especially when there is a competition between two or more creatures. The idea is that rather than evolving a robot against a fixed fitness function, two robots are used instead, with one evolving against the other. For example, Hillis evolved efficient number sorting algorithms by co-evolving a population of candidate algorithms against a population of numbers to be sorted [5]. As the sorting algorithms got improved, so are the population of numbers to be sorted, evolved to present tougher challenges to the algorithms.

3.2 Virtual Ecologies

Here we can concurrently simulate no of robots of same type having same fitness function against virtual world parameters like survival time, ability to complete the task and so on based on natural selection. In order to identify complex and non complex behaviors some steps have been proposed [7]. Now single physical simulation performed concurrently in parallel mode to test against other robots.

3.3 Lifetime Learning

Virtual robot controller gets improved by evolution, but they do not actually adapt whilst an individual robot is being simulated. Some recent results from evolutionary robotics suggest that combining an evolutionary algorithm with the ability of the controllers' to adapt or learn over an individual robot's lifetime can lead to improved robustness and complexity of behaviors compared to evolution by itself [4]. Giving an individual robot the ability to adapt and learn during its lifetime effectively smoothes the search space over which evolution is happening, thereby helping the process to progress. It is reasonable to assume that adding these sorts of abilities to our artificial robot will improve its ability to evolve complex behaviors just as it has done in evolutionary robotics.

3.4 Behavioral Primitives

The evolution of complex behaviors can in general be evolved by task decomposition. Rather than trying to evolve complex behaviors, another approach is to evolve a collection of primitive behaviors, and then use other, non-evolutionary techniques for combining these primitive into more complicated sequences. The task of programming a robot using this approach is like commanding it perform an action in the virtual world.

3.5 User Guided Evolution

Another alternative to supplying a fixed fitness function to the genetic algorithm is to present the user with a variety of robots from the evolving population at various intervals, and allow them to select their favorite prototypes to be used as the foundation of the next generation. The user may select it under any criteria, and can therefore guide the path of evolution according to their own preferences without having to formally instruct the individual robots.

4 Moving Robot Design

Simulation and optimization of digital robot were attempted by several researchers. In this case we used Framsticks software to simulate and evolve digital robots.

4.1 Framsticks

Framsticks is 3D simulation software for agents and controllers. It allows using user-defined experiment setups, fitness functions, and neurons (sensor network) and is suitable for testing various research hypotheses where fast 3D simulation and evolutionary optimization is required. The physical structure of Framsticks agents is made of parts (material points) and joints. The control system is made of neurons (including sensors and actuators) and their connections. Framsticks supports multiple genetic representations and operators and it ranges from simple and direct descriptions of agents to the ones encoding developmental process in genotypes. Further possibilities are like performing simulation of several real life creatures and storing in a library to use later.

4.2 Autonomous Moving Robot

Khepera robot having 2 wheels (left, right) and 10 sensors in all directions to pick the data from the nearby places and to decide the direction of navigation with out hitting obstacles or reaching a target in the virtual world. Sensor data will be passed to the brain or NN where decision will be made to select the direction based on threshold values of each neuron. Khepera robot is having a neural network associated with it to make decisions based upon real time parameters that are picked up from the environment. Here each node in represents a neuron (i.e.) two neurons for each wheel and a neuron for each sensor as well. Other parameters like ambience, surface type, obstacle sensors all to have its effect over the decision which would be taken by the robot before making a movement as shown in Figure 2 and Figure 3.

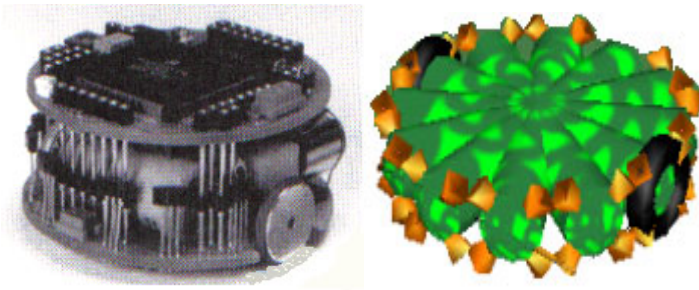


Fig. 1. (a) K-Team’s Khepera Robot (b) Framsticks Khepera Robot

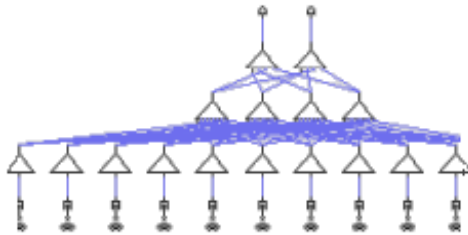


Fig. 2. Neural network of Khepera Robot

4.3 Fitness Function

In order to analyze the behavioral parameters of Khepera robot we need to define a fitness function that can monitor the properties like time taken to perform an action, no of rotations performed by each wheels.

$$F(t) = \{ \sum Si-Th \Rightarrow (L_w, R_w) \Rightarrow r \} \tag{1}$$

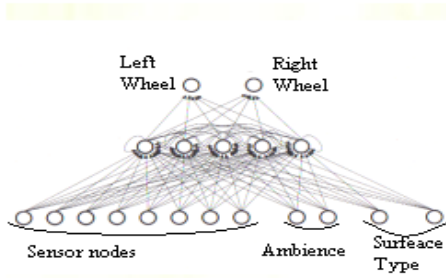


Fig. 3. Khepera Neural net connection pattern in Framsticks

In the above fitness function S_i represents signal from all sensor to be compared against T_h (threshold value) and movement made by left (L_w) and right (R_w) sided wheels in one unit time or to reach a target from the current location where as r denotes no of rotations performed to reach target.

4.4 Dynamic Simulation

Now we are going to design an autonomous moving robot by first creating the same in framsticks' virtual world. After that we will apply several experimental parameters to test the stability of robot in different environments available in framsticks (e.g. flat land, bumpy land) as per the AI techniques that are proposed earlier and both artificial and real robots are shown in Figure 1.

Khepera robot's kinematics structure is defined using genome editor exclusively available for framsticks and here we use f1 format with the following genomic sequence as it describes the robot in terms of small segments in framsticks virtual world.

`(rrX(IX(IISSEEX[T:1]),lmXMMMMMEEX[[1:2,-1:-3]rrSEEX[T:-0.407](SSISSLIEEX,,SSISSLIEEX,))`

After creating this structure we have simulated the above the digital creature in the framsticks virtual world with several environmental conditions with different set of parametric values. Simulation was conducted in a system with configuration of Intel Pentium Dual Core E2140 @ 2.8 GHz and 2 GB RAM in framsticks 3.0.

5 Experiments and Results

We have performed the simulation by applying the fitness function for different type of actions with the Khepera robot. Walking fitness function will be calculated from the distance between the robot wheels and the surface by setting default unit value 1. if this condition is not satisfied or violated then a penalty will be applied to over all fitness of the robot. In subsequent tests the movement of the wheels based on no of rotations from one location to the other in a stipulated time limit was tested. Then the above sequence was repeated with fast movement by reducing the time limit and again testing the fitness functional parameters of the robot.

Simulation was divided in to discrete steps translating the simulated system from starting time to finishing time $T_s - T_f + C$ where C is coarseness of simulation. Mean while other simulation parameters such as pressure applied to the physical parts of the robot was also monitored to predict the overall performance of the robot i.e. force and torque. During simulation we have monitored several parameters that are major concern in designing the robot like position, velocity, and acceleration and so on.

The aim of our experiment is to develop real time AI controllers which will reduce the time involved in building real time robots to test and also to provide many possible prototypes with limited resources. Especially in the research cases simulation parameters can be used to reduce the complexities in designing robots.

5.1 Khepera Robot in a Flat Surface

To analyze the behavior of digital robot in the virtual world we have selected a flat land surface where there are no bumps and disturbances. During this test we have started with the fitness value as 0 and after 30 minutes of simulation to reach the target in a flat land with and without obstacles we achieved the fitness function value of 0.7 and the average fitness function for reaching the target by relocating the robot in other corners of the virtual world was improving to 0.9. Further if we increase the no of simulation steps above 1,000 certainly we can reach the maximum fitness level. Failure rate for sensors and joints were in the region of 2 out of 25 and out of 25 for with and without obstacles respectively.

5.2 Khepera Robot in a Bumpy Surface

In bumpy surface we have performed movement test for the Khepera robot. Since there is a penalty for each move if it does not satisfy the distance between wheels and surface. Here we had some diversified results. First with out obstacles we got a fitness value of 0.4 and by increasing them slowly over 2000 steps in the same time limit of 30 minutes it has reached the maximum fitness value of 0.6. Performance of the robot suddenly changes due to the influence of change in environmental parameters.

Note: - Test results shown in the table I is for target reaching test of about 25 times in two surfaces and they are only average values under a simulation time of maximum 30 minutes.

Table 1. Comparison of Virtual Robot's Performance in different environments

S. No	Name of the Parameter	Flat Surface		Bumpy Surface		Combined	
		Without obstacle	With Obstacle	Without obstacle	With Obstacle	Without obstacle	With Obstacle
1	Fitness value	0 to 0.9	0 to 0.7	0 to 0.6	0 to 0.4	0.6	0.3
2	No of Rotations by Left wheel	13	27	38	42	27	33
3	No of Rotations by Right wheel	14	23	36	44	39	42
4	Sensor Performance	80%	65%	58%	49%	58%	49%
5	Failiure Rate	2/25	3/25	4/25	4/25	4/25	4/25

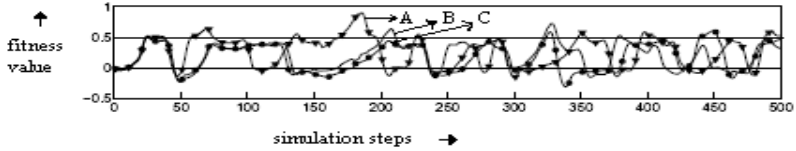


Fig. 4. Fitness graphs for Flat, Bumpy and Combined surfaces

Figure 4 shows the effect of sensors in joints and angles of the robot with respect to the user defined fitness function including penalty value for keeping low distance between surface and robot's bottom section to avoid physical component failures in the virtual world A)flat B)Bumpy C)Combined. Using the above data obtained by from simulation using Framsticks. Simulation data will help selecting the right sort of hardware components like motors used in joints, sensors, control programs and even walking strategies also. Several parameters like acceleration, force, and walking strategies can be used for physical robot design which reduces considerable time. Further this could be extended to other applications like surgical robots, industrial robots, land exploring robots also.

6 Conclusions

With the help of artificial simulation, we were able to capture the body shape of Khepera robot and to identify complex behaviors of the same with different fitness options. This method helps capturing the finest details / movements of the robot in a set of artificial environments. Methods suggested in this paper vindicate that effective construction of digital robot by evolving them in an artificial environment provides sufficient information to design the physical controllers in order to emulate them in a successful way especially in games and automation industries. It also avoid the risk involved in real time robot design and provides plenty of minute information to pick individual components and by concentrating on small data items we can design robot of finest quality with considerably less failure rate.

References

1. Arnold, D.: Evolution of Legged Locomotion. MSc thesis, School of Computing Science, Simon Fraser University (1997)
2. Bongard, J.C., Paul, C.: Investigating Morphological Symmetry and Locomotive Efficiency Using Virtual Embodied Evolution. In: Meyer, J.-A., et al. (eds.) From Animals to Animats: The Sixth International Conference on the Simulation of Adaptive Behavior. MIT Press, Cambridge (2000)
3. Dorigo, M., Colombetti, M.: Robot Shaping: An Experiment in Behavior Engineering. MIT Press, Cambridge (1997)
4. Floreano, D., Urzelai, J.: Evolution of Neural Controllers with Adaptive Synapses and Compact Genetic Encoding. In: Floreano, D., et al. (eds.) Advances in Artificial Life: Proceedings of the Fifth European Conference on Artificial Life, pp. 183–194. Springer, Heidelberg (1999)

5. Hillis, W.D.: Co-evolving parasites improves simulated evolution as an optimization technique. In: Langton, C.G., et al. (eds.) *Artificial Life II*, pp. 313–384. Addison-Wesley, Reading (1991)
6. Komosinski, M., Ulatowski, S.: Framsticks: Towards a Simulation of a Nature-Like World, Creatures and Evolution. In: Floreano, D., et al. (eds.) *Advances in Artificial Life: Proceedings of the Fifth European Conference on Artificial Life*, pp. 261–265. Springer, Heidelberg (1999)
7. Van de Panne, M., Fiume, E.: Sensor- Actuator Networks. In: *Proceedings SIGGRAPH 1993, Computer Graphics*, pp. 335–342. ACM SIGGRAPH (1993)

Analysis and Comparison of Position Based Routing Protocol in Mobile Adhoc Network

Vimal Upadhyay, Gori Shankar, Amritpal, and Jai Balwan

Department .of CS/EC ,St. Margaret Engineering College, Neemrana Alwar Rajasthan
vimalupadhyay2002@gmail.com
Contact Details: 09785132144

Abstract. This Paper introduces the concept of ad hoc network and its applications. It also presents a brief review of conventional routing protocols and various routing approaches proposed for mobile ad hoc network. A brief description of various parameters used for performance evaluation of routing algorithms is also presented. An ad hoc wireless network is “a collection of two or more devices equipped with wireless communications and network capability” that creates a network “on the fly” without any network infrastructure. They offer quick and easy network deployment in situations where it is not possible otherwise. That’s why the idea of ad hoc networking is sometimes also called infrastructure less networking. Mobile Ad-hoc networks are self-organizing and self-configuring multi-hop wireless networks where, the structure of the network changes dynamically. This paper is intended to bring out the details of mainly two Position Based Routing Protocol named as LAR, DREAM. In position based routing the position (location) plays an important role in routing decisions. The position of a node can be made available by using GPS receivers or by evaluating relative positions using signal strength .Position based routing does not require knowledge of the entire network to find a route from source to destination Position based routing require that information about the physical position of the participating nodes be available and describing various parameters for LAR ,DREAM used by various authors and Proposed parameters for evaluating their performance.

1 Introduction

An ad-hoc network is a collection of wireless mobile hosts forming a temporary network without the assistance of any stand-alone infrastructure or centralized administration. They offer quick and easy network deployment in situations where it is not possible otherwise. That’s why the idea of ad hoc networking is sometimes also called infrastructure less networking. Ad hoc networks are useful for the applications where central or fixed infrastructure is not available. Mobile Ad-hoc networks are self-organizing and self-configuring multi-hop wireless networks where, the structure of the network changes dynamically. This is mainly due to the mobility of the nodes Nodes in these networks utilize the same random access wireless channel, cooperating in a friendly manner to engaging themselves in multi-hop forwarding. The nodes in

the network not only act as hosts but also as routers that route data to/from other nodes in network Figure 1 represents a MANET of 3 nodes. Node 2 can directly communicate with node1 and node 3, but any communication between Nodes 1 and 3 must be routed through node 2.

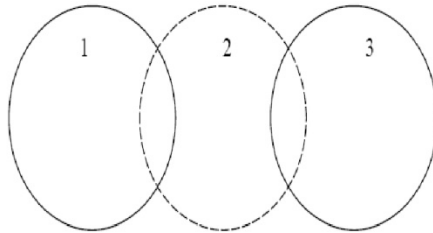


Fig. 1. Example of simple MANET of 3-nodes

Advantages of Mobile Ad-Hoc Networks

There are several advantages of using mobile ad hoc network.

1. Setting up a wireless system is easy and fast and it eliminates the need for pulling out the cables through walls and ceilings.
2. Network can be extended to places, which cannot be wired.
3. Multiple paths increase reliability.
4. Wireless network offers more flexibility and adapt easily to changes in the configuration of the network.

Limitations of Mobile Ad Hoc Networks

- Asymmetric links
- Routing Overhead
- Interference
- Dynamic Topology

2 Classification of Routing Protocol in Manet's

Classification of routing protocol in MANET depends on routing strategy and network structure. According to the routing strategy the routing protocols can be categorized as Table-driven and Source initiated, while depending on the network structure these are classified as flat routing, hierarchical routing and geographic position assisted routing. Both the Table-driven and source initiated protocols come under the Flat routing.

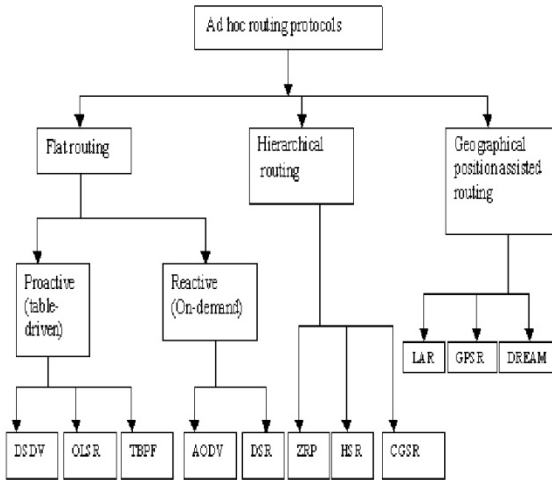


Fig. 2. Classification of Routing Protocols in MANET

3 Position Based Routing Protocol

Position based routing algorithms eliminate some of the limitations of topology-based routing by using additional information. They require that information about the physical position (that is their coordinates in two or three-dimensional space) of the participating nodes be available. The position of nodes may be available directly by communicating with a satellite, using GPS (Global Positioning System), if nodes are equipped with a small low power GPS receiver or some other type of positioning service. A location service is used by the sender of a packet to determine the position of the destination and to include it in the packet's destination address. Examples of position based routing algorithms are Location Aided Routing (LAR), Distance Routing Effect Algorithm for Mobility (DREAM), Greedy Perimeter Stateless Routing (GPSR), Compass Routing (referred as DIR), Geographic Distance Routing (GEDIR), Most Forward progress within Radius (MFR), Greedy Face Greedy (GFG), Grid Location Service (GLS), etc.

Advantages of Position Based Routing Are:

1. In wireless networks, the presence of geographically nearby nodes determines the existence of links.
2. A node can save battery power by adjusting the transmission power based on the availability of distance information of the destination. By saving battery power a node can increase the routing tasks in a network.
3. These algorithms do not require knowledge of the entire network topology.
4. Transmission of message from source can be restricted into a selected region based on the direction of the destination.

3.1 Location Aided Routing Protocol (LAR)

Location Aided Routing (LAR) protocol is a reactive protocol in which routes to destinations are determined only when explicitly needed to route packets. It uses location information to limit the route query flooding area. Every mobile host node is assumed to know its location and the global time, which can be provided by a Global Positioning System (GPS) or location services. The Location - Aided Routing Protocol uses location information to reduce routing overhead of the ad-hoc network. Normally the LAR protocol uses the GPS to get these location information. With the availability of GPS, the mobile hosts knows there physical location.

Expected Zone and Request Zone

Expected zone

First, we consider that the node S (source) needs to find a way to node D (destination). Node S knows that D was at position L. Then, the "expected zone" of node D, from the viewpoint of node S, is the region that node S expects to contain node D. When node S knows, that node D travels with a several speed, node S considers this speed to determine the expected zone of D.

When node S has no information about the position of D, the entire region of the ad-hoc network is assumed to be the expected zone. Then the algorithm reduces to the

Basic flooding algorithms. In general we can say, as more as the node S knows about D, as smaller can be considered the expected zone.

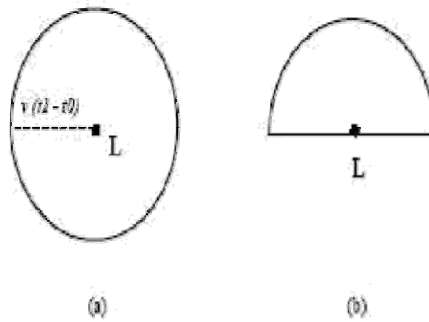


Fig. 3. (a) The circular expected zone of node D. (b) When node S knows that D is moving north, then the circular expected zone can be reduced to a semi-circle.

Request Zone

Again, we consider that node S needs to find a path to node D. Node S defines a Request Zone like in figure 3 (a). After that S sends a route request like in the normal flooding algorithm. With the difference, that a node only forwards this route request when it is in the Request Zone. But there are two different reasons, that regions outside the request zone have to be include in the Request Zone.

1) Node S is not in the Expected Zone of D, then this Expected Zone has to be enlarged to the Request Zone like in figure 4 (a).

But now we ask us if this Request Zone of figure 4 (a) is a good Request Zone We see in figure4 (b) that all the nodes between S and D are outside of the Request Zone. So it is not guaranteed, that a path between S and D can be found. LAR allows to expand the Request Zone, so that it easier to find a path. But we have to consider, when we increase the Expected Zone like in figure 4 (c), the route discovery overhead will also increase.

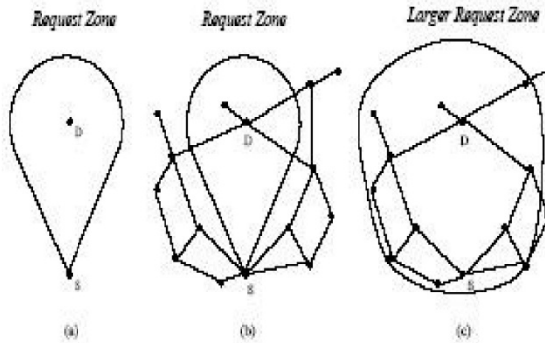


Fig. 4.

3.2 Distance Routing Effect Algorithm for Mobility (DREAM)

DREAM is a proactive protocol. It is a location-based algorithm. Each node MN in this proactive protocol maintains a location Table for all other nodes in the Ad hoc network. The frequency of distribution of location information depends upon the principles of Distance Effect and Mobility rate. Each node can sense these two principles with the help of information obtained from GPS. In general, for the first principal i.e. Distance Effect ,as shown in the Figure 5, the greater the distance separating two nodes the slower they appear to be moving with respect to each other. With the mobility rate the faster a node moves the more frequently it needs to advertise its new location. Using the location information obtained from GPS each node can realize the two principles in routing.

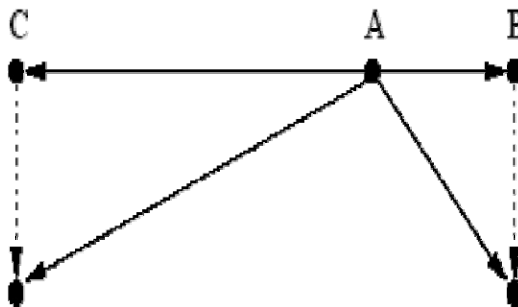


Fig. 5. Distance Effect

In other words, **DREAM** can be define as :

1. Basically a unicast algorithm, for mobile ad hoc networks
2. Proactive protocol
3. Routing tables stored at each node stores location information
4. Each node transmits control messages bearing its current location to all the other nodes. The frequency with which these control messages are transmitted is determined by “distance effect” and “mobility rate”

Advantages of DREAM are:

- Bandwidth & energy efficient
- Loop-free
- Robustness
- Adaptive to mobility
- Low average end-to-end delay

4 Performance Evaluation

Giordano, Stojmenovic and Blazevic [7] have suggested the following qualitative properties and quantitative metrics for measuring the performance of routing algorithms in ad hoc networks. These properties and metrics are independent of any routing protocol.

4.1 Qualitative Properties

- Distributed operation: A routing algorithm should be distributive in nature. Each node participating in routing should take independent routing decisions.
- Loop-freedom: A mechanism to avoid the circulation of a small fraction of packets around in the network for arbitrary time period.
- Demand based operation: Instead of assuming uniform traffic distribution within the network, the routing algorithm should adapt to the varying demand pattern. It will utilize the network resources effectively.
- Sleep period operation: Some nodes in a network may stop transmitting and/or receiving messages for an arbitrary time period to save battery power or for some other reasons. A routing protocol should be able to accommodate such sleep periods without overly adverse consequences.

4.2 Quantitative Properties

- End to end data throughput: This is a measure of the rate at which data can be sent through a network between source and destination.
- End to end data delay: This is also referred to as latency. It is the time needed to deliver a message from a source to destination.
- Average number of data bits transmitted per data bit delivered: This measures the efficiency of delivering data within a network.

- Average number of control bits transmitted per data bit delivered: This measures the efficiency of a protocol in expending control overheads to deliver data packets. Anything (such as control packets, header of data packets, etc.) that is not data should be included in control overheads.

5 Proposed Solution for Performance Evaluation

- Percentage of successful deliveries: It is calculated as the sum of the number of messages delivered successfully to destination divided by the total number of messages sent.
- Average minimum hop counts: It is the sum of hop counts of all successful deliveries divided by total number of successful deliveries. In case there is more than one successful delivery for a particular source-destination pair, only those successful deliveries with minimum hop count are considered in the sum of hop counts.
- Percentage of nodes connected within a circle: It is calculated as the total numbers of the mobile nodes which receives message from a particular node.

6 Conclusions

This work is an attempt to simulate position based algorithms. The performance comparison of various algorithms may help in choosing routing algorithm for a given network. This may assist in designing an ad hoc network by determining right values of parameters in a given situation so that the performance of the network does not degrade.

References

1. Lang, D.: A Comprehensive overview about selected Ad-Hoc Networking Routing Protocols, Technical Report, Department of Computer Science, Teclmische University (March 14, 2003)
2. Ramakrishanan, K.: An Improved Model for the Dynamic Effect Algorithm for Mobility protocol, PhD thesis, Department of Science, University of Waterloo (2004)
3. Ko, Y.-B., Vaidhya, N.H.: Location-Aided Routing in Mobile Adhoc Networks. In: Proc. Of MOBICOM 1998 (1998)
4. Basagni, S., Chlamtac, I., Syrotiuk, V.R., Woodward, B.A.: A Distance Routing Effect Algorithm for Mobility (DREAM). In: Proc. ACM/IEEE Intl. Conference on Mobile Computing and Networking (MOBICOM), Dallas, TX, pp. 76–84 (October 1998)
5. Su, W., Lee, S.-J., Gerla, M.: Mobility Prediction and Routing in Adhoc Wireless Networks, Department of Computer Science, University of California
(wsu, sjlee, gerla)@cs.ucla.edu

6. Luo, X.: Predictive Methods for Location Services in Mobile Ad hoc Networks, Technical Report, Department of Mathematical and Computer Science (2004)
7. Karp, B., Kung, H.T.: GPSR: Greedy Perimeter Stateless Routing for Wireless Networks. In: MobiCom 2000, Conference of the IEEE Computer and Communication Societies (INFOCOM 2002) (2000)
8. The Network Simulator – ns-2, <http://www.isi.edu/nsnam/ns>
9. Mobile Ad-hoc Networks (MANET), <http://www.ietf.org/html.charters/manetcharter.html>
10. <http://www.isi.edu/nsnam/tutorial> -Marc Greis Tutorial for ns-2

Author Index

- Abbiramy, V.S. 220
Adinarayana, G. 23
Agarwal, Harshit 75
Aghila, G. 257
Alagappan, M. 146
Amritpal 468
Anandhakumar, P. 86
Anburajan, M. 383
Arun, R. 146
Ashish, B.K. 277
- Badari Nath, D. 405
Balwan, Jai 468
Bardhan, Prajnat 229
Basheera, Shaik 185
- Chakraborty, Aruna 229
Chandra Murty, Patnala S.R. 165
Chellapa, Murugesan 393
Chowdhury, Shalini 173
- David, Mahibha 44
Deepika, C. Lakshmi 146
Dey, Debika 173
Dhinaharan, A.L. 378
- Elumalai, Giridara Varma 393
Eslamnour, Behdis 438
Evangelin Geetha, D. 346
- Felci Rajam, I. 11
- Gajjar, Trusha 55
Gambhir, Deepak 96
Geeta, R.B. 239
Gohil, Gunvantsinh 55
Gopi, Aswin Kumar 393
Goyani, Mahesh 55
- Jaganathan, P. 203
Jagannathan, S. 438
Jagatramka, Ankul 229
James, Rekha K. 417
Jeya Nachiabam, N.M. 128
- Jhansi Vazram, B. 336
Jindal, M.K. 268
- Kandaswamy, A. 146
Kantikiran, M. 23
Kokare, Manesh B. 108, 118
Kole, Dipak K. 229
Kumar, A. 277
Kumar, Munish 268
- Laha, Samanwita 173
Lakshmi Sirisha, B. 23
- Mamillapalli, Omkar 239
Markan, Chota 75
Meyyappan, T. 128
Mohankumar, N. 405
Mouli, Chandra 291
Murthy, J.V.R. 336
Murugan, D. 378
- Nachiappan, Alamelu 429
Nagamalai, Dhinaharann 366
Naganjaneyulu, P.V. 185
Natarajan, Sudha 249
Nirmala Devi, M. 405
- Padhy, P.K. 277
Patil, Pushpa B. 118
Poulose Jacob, K. 417
Prakash, D. Bhanu 185
Pramanick, Swarnali 229
Prasad Reddy, P.V.G.D. 239
Prasanna Venkatesan, V. 257
Prudhvi Raj, V. Naga 212
Pugalendhi, Ganesh Kumar 44
Puhan, Niladri B. 249
- Rajagopal, Karthikeyan 193
Rajalakshmi, K. 378
Raja Mohamed, S. 460
Rajani Kanth, K. 346
Rajesh, S. 212
Rajinikannan, M. 203
Rajpal, Navin 96

- Raju, K.V.S.V.N. 326
 Ramachandran, Sumalatha 393
 Ram Mohan Reddy, Ch. 346
 Raviraj, P. 460
 Ribeiro, Sérgio 449
 Roy, Sudipta 1

 Sampath, Sivaperumal 156
 Santhi, G. 429
 Saphthagirivasan, V. 383
 Sarangdevat, S.S. 355
 Saravanakumar, S.G. 315
 Sarker, Subhojit 173
 Sasi, Sreela 417
 Scaria, Arun 405
 Selvarani, R. 366
 Sen, Asoke Kr. 1
 Shankar, Gori 468
 Sharma, Ritu K. 268, 305
 Sharma, Sheena 75
 Shirdhonkar, M.S. 108
 Singaravelan, S. 378
 Singh, Sanjay 291
 Singh, Vipula 96
 Sinha, Nidul 1
 Sinha, Sanjana 229
 Sood, Manu 305
 Sreenivasa Reddy, E. 165
 Srinivas, P. 212
 Srinivasa, K.G. 346
 Sri Rama Krishna, K. 23
 Suchithra, R. 366

 Sudheer, T. 165
 Suguna, R. 86
 Suhas Hegde, A. 249
 Suiama, Danilo 449
 Sultana, Najiya 355
 Suresh Kumar, T.V. 346
 Suresh Varma, P. 65

 Tamilarasi, A. 220
 Tamililakkiya, V. 34
 Teraiya, Rekha 55
 Thamarai, S.M. 128
 Thiyagarajan, P. 257
 Thomas, Deepa Mary 136
 Totad, Shasikumar G. 239

 Upadhyay, Vimal 468

 Vadaparthi, Nagesh 65
 Vaidyanathan, Sundarapandian 156,
 193
 Valli, S. 11
 Valli Kumari, V. 326, 336
 Vani, K. 34
 Venkata Ramana, K. 326
 Victoire, Aruldoss Albert 44
 Vijaykumar, S. 315

 Wassim Ferose, H. 146

 Yarramalle, Srinivas 65

 Zawodniok, Maciej 438