

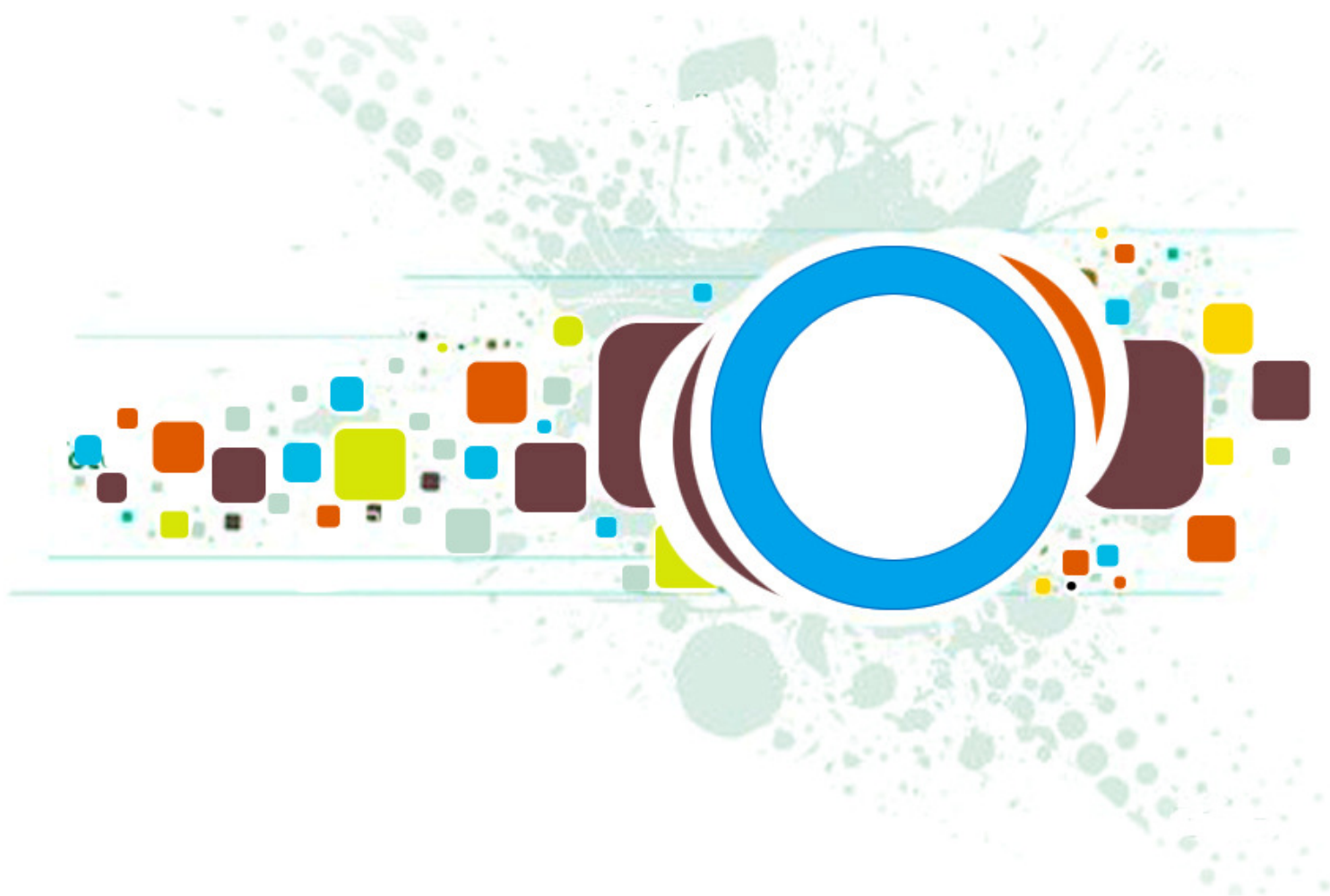
Volume 5 ▪ Issue 2 ▪ May 2011

Editor-in-Chief
Professor Hu, Yu-Chen

INTERNATIONAL JOURNAL OF
IMAGE PROCESSING (IJIP)

ISSN : 1985-2304

Publication Frequency: 6 Issues Per Year



CSC PUBLISHERS
<http://www.cscjournals.org>

INTERNATIONAL JOURNAL OF IMAGE PROCESSING (IJIP)

VOLUME 5, ISSUE 2, 2011

**EDITED BY
DR. NABEEL TAHIR**

ISSN (Online): 1985-2304

International Journal of Image Processing (IJIP) is published both in traditional paper form and in Internet. This journal is published at the website <http://www.cscjournals.org>, maintained by Computer Science Journals (CSC Journals), Malaysia.

IJIP Journal is a part of CSC Publishers

Computer Science Journals

<http://www.cscjournals.org>

INTERNATIONAL JOURNAL OF IMAGE PROCESSING (IJIP)

Book: Volume 5, Issue 2, May 2011

Publishing Date: 31-05-2011

ISSN (Online): 1985-2304

This work is subjected 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 of parts thereof is permitted only under the provision of the copyright law 1965, in its current version, and permission of use must always be obtained from CSC Publishers.

IJIP Journal is a part of CSC Publishers

<http://www.cscjournals.org>

© IJIP Journal

Published in Malaysia

Typesetting: Camera-ready by author, data conversion by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2011

EDITORIAL PREFACE

The International Journal of Image Processing (IJIP) is an effective medium for interchange of high quality theoretical and applied research in the Image Processing domain from theoretical research to application development. This is the forth issue of volume four of IJIP. The Journal is published bi-monthly, with papers being peer reviewed to high international standards. IJIP emphasizes on efficient and effective image technologies, and provides a central for a deeper understanding in the discipline by encouraging the quantitative comparison and performance evaluation of the emerging components of image processing. IJIP comprehensively cover the system, processing and application aspects of image processing. Some of the important topics are architecture of imaging and vision systems, chemical and spectral sensitization, coding and transmission, generation and display, image processing: coding analysis and recognition, photopolymers, visual inspection etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 5, 2011, IJIP appears in more focused issues. Besides normal publications, IJIP intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

IJIP give an opportunity to scientists, researchers, engineers and vendors from different disciplines of image processing to share the ideas, identify problems, investigate relevant issues, share common interests, explore new approaches, and initiate possible collaborative research and system development. This journal is helpful for the researchers and R&D engineers, scientists all those persons who are involve in image processing in any shape.

Highly professional scholars give their efforts, valuable time, expertise and motivation to IJIP as Editorial board members. All submissions are evaluated by the International Editorial Board. The International Editorial Board ensures that significant developments in image processing from around the world are reflected in the IJIP publications.

IJIP editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJIP. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. IJIP provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members

International Journal of Image Processing (IJIP)

EDITORIAL BOARD

EDITOR-in-CHIEF (EiC)

Professor Hu, Yu-Chen
Providence University (Taiwan)

ASSOCIATE EDITORS (AEiCs)

Professor. Khan M. Iftekaruddin
University of Memphis
United States of America

Dr. Jane(Jia) You
The Hong Kong Polytechnic University
China

Professor. Davide La Torre
University of Milan
Italy

Professor. Ryszard S. Choras
University of Technology & Life Sciences
Poland

Dr. Huiyu Zhou
Queen's University Belfast
United Kingdom

Professor Yen-Wei Chen
Ritsumeikan University
Japan

EDITORIAL BOARD MEMBERS (EBMs)

Assistant Professor. M. Emre Celebi
Louisiana State University in Shreveport
United States of America

Professor. Herb Kunze
University of Guelph
Canada

Professor Karray Fakhreddine
University of Waterloo
United States of America

Assistant Professor. Yufang Tracy Bao
Fayetteville State University
North Carolina

Dr. C. Saravanan

National Institute of Technology, Durgapur West Benga
India

Dr. Ghassan Adnan Hamid Al-Kindi

Sohar University
Oman

Dr. Cho Siu Yeung David

Nanyang Technological University
Singapore

Dr. E. Sreenivasa Reddy

Vasireddy Venkatadri Institute of Technology
India

Dr. Khalid Mohamed Hosny

Zagazig University
Egypt

Dr. Gerald Schaefer

Loughborough University
United Kingdom

Dr. Chin-Feng Lee

Chaoyang University of Technology
Taiwan

Associate Professor. Wang, Xiao-Nian

Tong Ji University
China

Professor. Yongping Zhang

Ningbo University of Technology
China

Professor Santhosh.P.Mathew

Mahatma Gandhi University
India

TABLE OF CONTENTS

Volume 5, Issue 2, May 2011

Pages

- 109 - 118 Recognition of Tifinaghe Characters Using a Multilayer Neural Network
Rachid El Yachi, Mohamed Fakir, Belaid Bouikhalene
- 119 - 129 Side-Match Vector Quantizers Using Neural Network Based Variance Predictor for Image Coding
Shuangteng Zhang
- 130 - 144 A Spectral Domain Dominant Feature Extraction Algorithm for Palm-print Recognition
Hafiz Imtiaz, Shaikh Anowarul Fattah
- 145 - 165 Quaternion Based Omnidirectional Machine Condition Monitoring System
Wai Kit Wong, Chu Kiong Loo, Way Soong Lim
- 166 - 176 Automatic Threshold based Liver Lesion Segmentation in Abdominal 2D-CT Images
Asmita A Moghe, Jyoti Singhai, S.C Shrivastava
- 177 – 184 Segmentation Based Multilevel Wide Band Compression for SAR Images Using Coiflet Wavelet
Parathasarthy Subashini , M. Krishnaveni
- 185-198 A Parametric Approach to Gait Signature Extraction for Human Motion Identification
Mohamed Rafi, Md. Ekramul Hamid, Mohamed Samiulla Khan, R.S.D Wahidabanu
- 199-219 An Experimental Study into Objective Quality Assessment of Watermarked Images
Anurag Mishra, Aruna Jain, Manish Narwaria, Charu Agarwal

220-235 Fully Automatic Method for 3D T1-Weighted Brain Magnetic Resonance Images Segmentation
Bouchaib Cherradi, Omar Bouattane, Mohamed Yousfi, Abdelhadi Raihani

Recognition of Tifinaghe Characters Using a Multilayer Neural Network

Rachid EL Ayachi

*Faculty of Sciences and Techniques/
Computer Sciences Department
Sultan Moulay Slimane University
Béni-Mellal, BP:523, Morocco*

rachidieea@yahoo.fr

Mohamed Fakir

*Faculty of Sciences and Techniques/
Computer Sciences Department
Sultan Moulay Slimane University
Béni-Mellal, BP:523, Morocco*

medfaki@yahoo.fr

Belaid Bouikhalene

*Faculty of Sciences and Techniques/ Information
processing and telecommunications teams
Sultan Moulay Slimane University
Béni-Mellal, BP:523, Morocco*

bbouikhalene@yahoo.fr

Abstract

In this paper, we present an off line Tifinaghe characters recognition system. Texts are scanned using a flatbed scanner. Digitized text are normalised, noise is reduced using a median filter, baseline skew is corrected by the use of the Hough transform, and text is segmented into line and lines into words. Features are extracted using the Walsh Transformation. Finally characters are recognized by a multilayer neural network.

Keywords: Tifinaghe Characters, Baseline Skew Correction, Segmentation, Walsh Transform, Hough Transform, Neural Network, Recognition.

1. INTRODUCTION

Optical Character Recognition (OCR) is one of the most successful applications of automatic pattern recognition. It is a very active field of research and development.

Several studies have been conducted on Latin, Arabic and Chinese characters [1, 2, 3, 4, 5, 6, 7, 8, 9]. However, for Tifinaghe characters system few works was done [13, 14, 15, 16].

Succession of operations in most digital image recognition system can be divided into three stages. First stage is a pre-processing including thresholding improving image quality, segmentation and son on. Second, features extraction for avoiding data abundance and reducing its dimension. Third stage is a classification. During this stage classes name is joint with unknown image by extracted features analyses and matching its representatives of the class, which the classifier has trained at a stage of training.

In this study a recognition system (Figure 1) for the recognition of Tifinaghe characters issued from an image scanner is presented. Initially, an image that contains Tifinaghe characters is normalized and segmented to produce a data base. Then, we applied the approach of Walsh Transform to extracted features which are used in the classification phase with a multilayer neural network.

The organisation of this paper is as follows. In section 2 characteristics of Tifinagh characters are given. In section 3 pre-processing process is described. Features extraction step is described in

section 4. Section 5 deals with the recognition step. Experimental results are given in section 6. Finally, this work is ended by a conclusion.

2. TIFINAGHE CHARACTERS

The Tifinaghe script is used by approximately 20 million people who speak varieties of languages commonly called Berber or Amazigh. The three main varieties in Morocco are known as Tarifite, Tamazighe, and Tachelhite. In Morocco, more than 50% of the population speaks Berber. In accordance with recent governmental decisions, the teaching of the Berber language, written in the Tifinaghe script, will be generalized and compulsory in Tifinaghe is an alphabetic writing system. It uses spaces to separate words and makes use of Western punctuation. The earliest variety of the Berber alphabet is Libyan. Two forms exist: a Western form and an Eastern form. The Western variety was used along the Mediterranean coast from Kabylia to Morocco and most probably to the Canary Islands. The Eastern variety, old Tifinaghe, is also called Libyan-Berber or old Tuareg. It contains signs not found in the Libyan variety and was used to transcribe Old Tuareg. A number of variants of Neo-Tifinaghe exist, the first of which was proposed in the 1960s by the Académie Berbère. That variant has spread in Morocco and in Algeria, especially in Kabylia. Other Neo-Tifinaghe systems are nearly identical to the Académie Berbère system. The encoding in the Tifinaghe block is based on the Neo-Tifinaghe systems. Historically, Berber texts did not have a fixed direction. Early inscriptions were written horizontally from left to right, from right to left, vertically (bottom to top, top to bottom); boustrophedon directionality was also known.

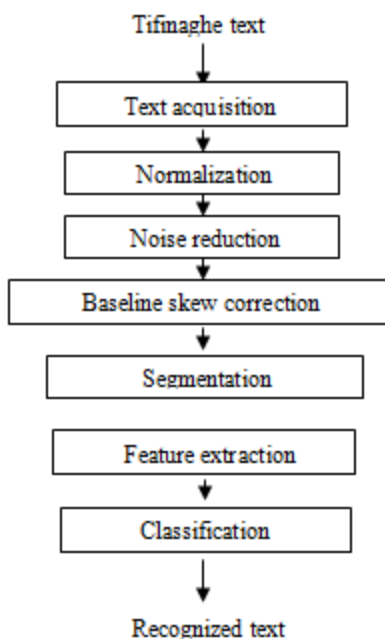


FIGURE 1: Tifinaghe recognized system.

Modern-day Berber script is most frequently written in horizontal lines from left to right; therefore the bidirectional class for Tifinaghe letters is specified as strong left to right. The encoding consists of four Tifinaghe character subsets: the basic set of the " Institut Royal de la Culture Amazighe (IRCAM) ", the extended IRCAM set, other Neo-Tifinaghe letters in use, and modern Tuareg letters. The first subset represents the set of characters chosen by IRCAM to unify the orthography of the different Moroccan modern day Berber dialects while using the historical Tifinaghe script. The alphabet Tifinaghe adopted by IRCAM [9] is composed of thirty-three characters representing consonants and vowels as shown in Table 1.

The table displays the Tifinaghe alphabet adopted by IRCAM. It is organized into two main sections. The top section shows the letters Σ, I, H, C, I with their corresponding Tifinaghe symbols. The bottom section shows a grid of 26 Tifinaghe symbols, each with a small Latin letter below it. The symbols are arranged in three rows: the first row has 5 symbols, the second row has 5 symbols, and the third row has 6 symbols. A legend is provided in the bottom right corner.

TABLE 1: Tifinaghe characters adopted by IRCAM.

3. PREPROCESSING

Pre-processing is the first part of Tifinaghe characters recognition system which covers four functions to produce a cleaned up version of the original image so that it can be used directly and efficiently by the feature extraction components of the OCR. These functions are: scanning the text and digitizing it into a digital image and cleaning it (by medium filter for example), converting the grey-scale image into binary image, normalizing the text, detecting and correcting Baseline Skew, and segmenting the text into lines and the lines into characters.

3.1 Normalization of the position

The position normalization is designed to eliminate unwanted areas and reduce the processing time. In this operation, firstly, we compute the horizontal and vertical histograms, secondly, we scan the horizontal histogram in two directions: from top to bottom and bottom to top respectively until the first meeting of black pixels, finally, we scan the vertical histogram in two directions: from left to right and right to left respectively until the first meeting of black pixels. After obtaining the positions of first black pixels, unwanted areas are eliminated in the image as shown in (Fig. 2).

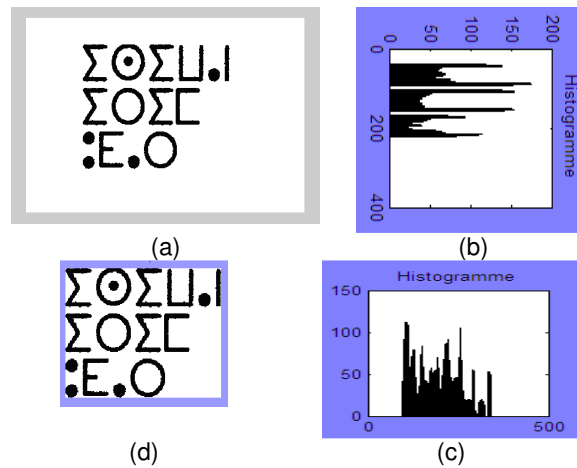


FIGURE 2: (a) Before normalization, (d) After normalization, (b) Horizontal histogram and (c) Vertical histogram

3.2 Baseline Skew Detection and Correction

A skew angle is the angle that the text lines of the document image make with the horizontal direction. The skew correction is necessary for the success of many OCR systems.

There are many methods to use for detecting skew angle, such as: the Trincklin method that uses the process of least squares to estimate skew angle, the Projection method is based on the calculation of horizontal histogram of the image, Hough transform, Fourier transform, Correlation lines, k-nearest neighbors. [17,18,19].

In this paper, the operation of skew correction is to estimate a skew angle θ_s using the Hough transform and to rotate the image by θ_s in the opposite direction, which gave the good results as showed in (Fig.3).

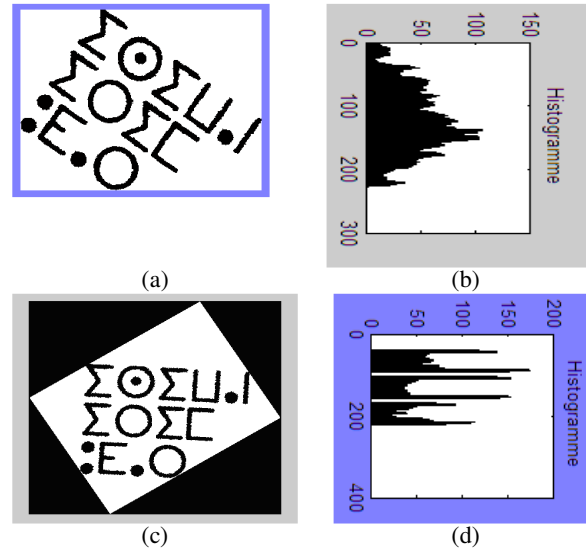


FIGURE 3: (a) Before correction, (c) After correction, (b) Horizontal histogram before correction, (d) Horizontal histogram after correction.

3.3 Segmentation

The last function to apply into pre-processing part is the segmentation; it is used to detect lines and characters in the image.

This method covers two steps: firstly, we use the horizontal histogram to detect lines; secondly, we use the vertical histogram to detect characters.

In the horizontal histogram, we browse from top to bottom until the first line containing at least one black pixel, the line is the beginning of the first line of text, then we continue traverse until a line that contains only white pixels, this line corresponds to the end of the first line of text. With the same way, we continue to detect other text lines.

In the vertical histogram, for each line of text, we browses from left to right until the first column containing at least one black pixel, this column is the beginning of the first character, then we continue traverse until a column that contains only white pixels, this column corresponds to the end of the first character. We continue detecting other characters of text with the same way.

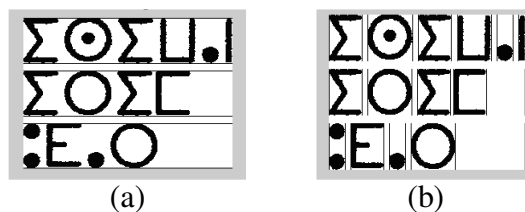


FIGURE 4: (a) lines segmentation, (b) Characters segmentation

4. FEATURES EXTRACTION

The second phase of Tifinaghe characters recognition system is Features extraction. Several methods can be used to compute the features: invariant momentsm Walsh transformation [20,21] etc.,.

In this recognition system, we use Walsh Transformation to extract features, because this method is independent to translation, rotation and scale change.

The Walsh transformation is given by:

$$W(u, v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y)g(x, y, u, v) \quad (2)$$

Where $f(x, y)$, is the intensity of the pixel with the coordinates (x, y) in the original binary image. The size of image f is $N*N$, and $u, v = 0, \dots, N - 1$, thus we compute N^2 Walsh transforms, $g(x, y, u, v)$ is the Kernel function given by the following form:

$$g(x, y, u, v) = (1/N) \prod_{i=0}^{n-1} (-1)^{b_i(x)b_{n-i+1}(u)+b_i(y)b_{n-i+1}(v)} \quad (3)$$

Where $b_i(x)$ is the i th bit in the binary expansion of x (it is equal either 0 or 1).

Table2 represents the seven first elements of the vector Walsh calculated for one character with his four transformations:

W_i	Σ	Σ	Σ	W
W_1	0	0	0	-0.0029
W_2	-0.0029	-0.0029	-0.0029	-0.0059
W_3	-0.0064	-0.0064	-0.0059	-0.0088
W_4	-0.0098	-0.0098	-0.0093	-0.0118
W_5	-0.0132	-0.0132	-0.0127	-0.0137
W_6	-0.0167	-0.0167	-0.0162	-0.0152
W_7	-0.0201	-0.0201	-0.0196	-0.0172

TABLE 2: Walsh Coefficients.

5. CHARACTER RECOGNITION

In the character recognition system, the recognition is the last phase which is used to identify the segmented character. Where we use the Neural Network approach for several reasons: the execution time is reduced and the principle of Neural Network is simple and effective.

In this phase a neural network is used [12], from an analogy with the biological neuron, is a processor that implements simple inputs and can connect with others to form a network that can achieve a relationship any entry-exit.

The Neural Network as shown in (Fig.5) represents an example of Neural Network multilayer which contains one hidden layer. It has:

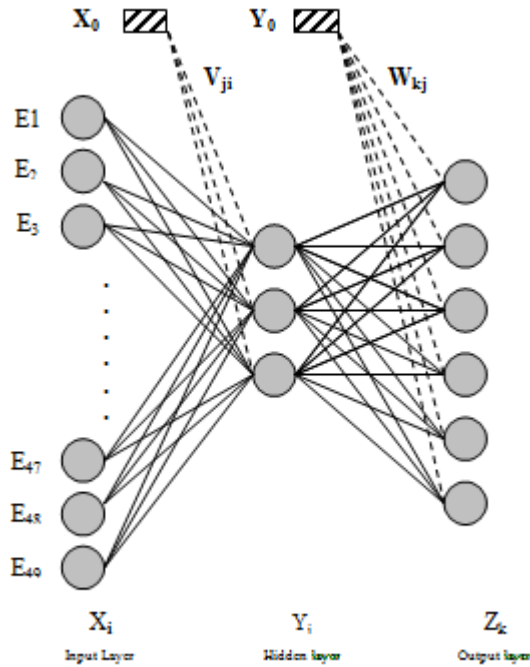


FIGURE 5: Neural Network

- An input layer of 49 (Walsh vector) inputs cells $E_i = X_i$ (the cells represents the inputs E_i of Network).
- A hidden layer of 3 activations Neural Y_j .
- An output layer of 6 activations Neural Z_k .
- 49×3 connections between input layer and hidden layer, each weighted by V_{ji} .
- 3×6 connections between hidden layer and output layer, each weighted by W_{kj} .
- X_0, Y_0 are initialled values (scalars).

The operation of Neural Network as shown in (Fig.5) contains five steps:

- **Step 1:** (Initializing weights of connexions), the weights are randomly selected.

- **Step 2:** (propagation of inputs)

The inputs E_i are presented to input layer: $X_i = E_i$.

We propagate to hidden layer:

$$Y_j = f \left(\sum_{i=1}^{49} X_i V_{ji} + X_0 \right) \quad (4)$$

After for hidden layer to output layer:

$$Z_k = f\left(\sum_{j=1}^3 Y_j W_{kj} + Y_0\right) \quad (5)$$

The values X_0 and Y_0 are means (scalars).

f is the activation function which is given by

$$f(x) = 1/(1 + \exp(-x)) \quad (6)$$

- **Step 3:** (Error back propagation)

For each example of applied learning base input of the network, we calculate the error at output layers, i.e. the difference between the desired output S_k and Z_k actual output:

$$E_k = Z_k (1 - Z_k)(S_k - Z_k) \quad (7)$$

We propagate this error on the hidden layer; the error of each neuron of the hidden layer is given by:

$$F_j = Y_j (1 - Y_j) \sum_{k=1}^6 W_{kj} . E_k \quad (8)$$

- **Step 4:** (Correction of connections weights)

We change the weights of connections:

- Between input layer and hidden layer:

$$\Delta V_{ji} = \eta . X_i . F_j \quad \text{And} \quad \Delta Y_0 = \eta . F_j \quad (9)$$

- Between hidden layer and output layer:

$$\Delta W_{kj} = \eta . Y_j . E_k \quad \text{And} \quad \Delta X_0 = \eta . E_k \quad (10)$$

Where η is the learning rate comprised between 0 and 1. This is experimentally determined ($\eta = 0.9$)

- **Step 5:** (Loop)

Loop in step tow to a criterion to define.

(Error threshold = 0.0001, Number of iterations = 50000)

After the learning of Network and the execution of Tifinaghe Characters Recognition System to recognize a Text, we use the Euclidian distance to identify the characters of Text.

$$d(t_k, o) = \left(\sum_{i=1}^6 (t_{ki} - o_i)^2\right)^{1/2} \quad (11)$$

Where, t_k is a desired output and o is the output of Network.

6. EXPERIMENTALS RESULTS

A Data Base used in this system contains 360 images which represents the Tifinaghe characters. All tests are applied on 158 characters.

Tests applied on several images gave the good results, which demonstrate the performance of the recognition system. Table 3 illustrated some recognized words.

Text to be recognized	recognition Results
ΣΟΣΣ	ΣΟΣΣ
⊙ΣΣ.	⊙ΣΣ.
Σ⊙ΣΠ.Ι	Σ⊙ΣΠ.Ι
∴E.O	∴E.O
+X.	+G.
Π.Ι	Π.Ι

TABLE 3: Examples of some words used for the test of the recognition method.

Number of hidden layer	Recognition rates	Error rates	Computing time
1	93.52%	6.48%	21.23s
2	86.71%	13.29%	28.65s
3	87.34%	12.66%	36.21s

TABLE 4: Recognition rates, Error rates and Computing times

A close inspection of Table 4 show that the recognition rate using one hidden layer is higher than those obtained by tow hidden layers, but error rates and computing time using two hidden layers are less than to those obtained by one hidden layer. Table5 illustrates the misrecognised characters. These error are caused by noise or rotation. The method has been implemented in Matlap software on a core (TM) Duo CPU T5870 @ 2.00 GHz

Noise	Rotation
○ ⊙	⊖ ⊗
	⊘ ⊙
	⊚ ⊛

TABLE 5: Misrecognised characters

7. CONCLUSION

The subject of the work developed in this paper is to achieve system recognition of Tifinaghe characters. This system consists of three phases applied on the input image: pre-processing features extraction and recognition. Pre-processing phase includes normalisation baseline skew correction and segmentation. The features extraction phase is used to compute the characters features using Walsh Transformation for the reasons of invariance to translation, rotation and scale change. In the recognition phase a multilayer neural network is used to classify characters. Experimental results showed that this method give good recognition rate in a final conclusion, neural network seems to be better than other techniques used for recognition

8. REFERENCES

[1] R. M. Bozinovic and S. N. Shihari, Off Line Cursive Script Word Recognition, IEEE Trans.Pattern Anal. Mach. Intell. PAMI 11, 1989, pp. 68- 83.

[2] M. K. Brown, pre-processing techniques for cursive word recognition, Pattern Recognition, Vol.13, N°5, pp: 447-451, 1983.

- [3] M. Fakir and C. Sodeyama, Recognition of Arabic printed Scripts by Dynamic Programing Matching Method, IECICE Trans. Inf & Syst, Vol. E76- D, No.2 Feb. [4]. M. FAKIR, Reconnaissance des Caractères Arabes Imprimés, Thesis, 2001, pp : 28-36, semlalia faculty of science Morocco, 2001.
pp: 31-37,1993.
- [4] N.Mezghani A.Cheret N.Mitiche, Bayes classification of online arabic characters by Gibbs modeling of class conditional densities, IEEE Trans PAMI Vol 30, issue 7, pp: 1121-1131, july 2008.
- [5] Asiri, A. and M.S. Khorsheed, 2005. Automatic processing of handwritten Arabic forms using neural networks. Proceeding of the World Academy of Science, Engineering and Technology, Aug. 2005, pp: 313-317.
- [6] Sarhan A.M. and O. Helalat, 2007. Arabic character recognition using ann networks and statistical analysis. Proceeding of European and Mediterranean Conference on Information Systems, June 24-26, Polytechnic University of Valencia, pp: 1-9.
- [7] Alnsour, A.J. and L.M. Alzoubady, 2006. Arabichandwritten characters recognized by neocognitron artificial neural network. J. Pure Appl. Sci., 3: 1- 17.
- [8] Hadjar, K. and R. Ingold, 2003. Arabic newspaper segmentation. Proceeding of 7th International Conference on Document Analysis and Recognition, Aug. 3-6, IEEE Computer Society, pp: 895-899.
- [9] Hamza, Ali A. Back Propagation Neural Network Arabic Characters Classification Module Utilizing Microsoft Word; Journal of Computer Science 4 (9): 744-751, 2008.
- [10] Ch. Choisy and A. Belaid, Cross- learning in analytic word recognition without segmentation, in Int. Journal on documentAnal. & Recognition IJDAR, 4(4): 281-289, 2002.
- [11] Y.X. Gu et al, Application of a multilayer tree in computer recognition of Chinese character, IEEE Trans. On PAMI-5, N°.1, pp: 83-89, 1983.
- [12] Standardisation Amazighe (Actes du organisé par le centre de l'Aménagement Linguistique, Rabat, 8-9, décembre 2003
- [13] M.Amrouch, Y. Es saady, A. Rachidi, M. El Yassa and D. Mammass, Printed Amazigh Character Recognition by a Hybrid Approach Based on Hidden Markov Models and the Hough Transform, 978-1-4244-3757-3/09/\$25.00 ©2009 IEEE.
- [14] Y. Es saady, M.Amrouch, A. Rachidi, M. El Yassa and D. Mammass, Reconnaissance de caractères Amazighes Imprimés par le Formalisme des Automates à états finis, SITCAM'09, Agadir-Maroc.
- [15] M. Fakir, B. Bouikhalene and K. Moro, Skeletonization Methods Evaluation for the Recognition of PrintedTifinaghe characters, SITCAM'09, Agadir-Maroc.
- [16] M. Blumenstein & C. K. Cheng & X. Y. Liu, 2002, New Preprocessing Techniques for Handwritten Word Recognition, in Proceedings of the Second IASTED International Conference on Visualization, Imaging and Image Processing (VIIP 2002), ACTA Press, Calgary, pp. 480-484.
- [17] D. J. Burr, A Normalizing Transform For Cursive Script Recognition,Proc. 6th Int. J. Conf. on Pattern Recognition Munich (1982), pp. 1027–1030

- [18] A. Sehad, L. Mezai, M.T. Laskri, M. Cheriet, Détection de l'inclinaison des documents arabes imprimés.
- [19] Attila Fazekas and Andras Hajdu Recognizing Type set Documents using Walsh , JCIT-CIT 9, 2-2001, 101-112.
- [20] Ibrahim S. I. Abuhaiba, Arabic Font Recognition Using Decision Trees Built From Common Words, JCIT-CIT 13, 3-2005, 211-223.

Side-Match Vector Quantizers Using Neural Network Based Variance Predictor for Image Coding

Shuangteng Zhang

*Department of Computer Science
Eastern Kentucky University
Richmond, KY 40475, U.S.A.*

shuangteng.zhang@eku.edu

Abstract

Side-match vector quantizer reduces bit-rates in image coding by using smaller-sized state codebooks generated from a master codebook through exploiting the correlations of neighboring vectors. This paper presents a new neural network based side-match vector quantization method for image coding. In this method, based on the variance of a vector which is predicted by a neural network, a subset of the codewords in the master codebook is selected for the side-matching to construct the state codebook for the encoding of the vector. This technique generates a lower encoding bit rate with a higher reconstructed image quality. Experimental results demonstrate that in terms of PSNR (Peak Signal-to-Noise Ratio) of the reconstructed images, the proposed method significantly outperforms the regular side-match vector quantizer, especially at lower coding bit-rates.

Keywords: Vector Quantization, Image Coding, Side Match, Neural Network.

1. INTRODUCTION

Image data compression is very important for various image and video processing applications which require reduced bit-rate/channel bandwidth. Such applications include digital television, video conferencing, telemedicine, multimedia, and remote sensing images from satellite and reconnaissance aircrafts, as well as the storage systems of multispectral image data from space programs, medical images, finger prints, and facial images.

There are many approaches to data compression. Among them, vector quantization (VQ) [1-3], which achieves data compression through mapping n-dimensional vectors onto a finite set of representative vectors called a codebook, is an important and heavily conducted research area. In VQ image compression [4], the images are usually partitioned into blocks with each block forming an n-dimensional vector. Each of these vectors is then coded as an index of its best matching vector in the codebook. The reconstructed images are obtained simply by selecting the corresponding codeword vectors from the codebook using the indices.

VQ image compression takes advantage of the correlation of the image pixels within a block (vector) for the coding bit-rate reduction and yields acceptable performance at low bit-rates. To further improve VQ's performance at low bit-rates, a side-match vector quantization method (SMVQ) [5] has been proposed. This method explores not only the redundancy within a vector but also the strong correlation between the neighboring vectors for high quality image coding at low bit-rates. In practice, SMVQ assumes the continuity of edges across neighboring image blocks' boundaries and generates a smaller-sized state codebook from the master codebook for each block's encoding. The selection of each state codebook is a subset of codewords in the master codebook which are the best matches of the upper and left blocks of the block to be encoded. In other words, SMVQ reduces the coding bit-rates by predicting the current block using its upper and left boundary blocks. Therefore, SMVQ's performance relies on the accuracy of the prediction. Inaccurate prediction, which occurs often when the block is within an edge area, may result in reduced quality of reconstructed images.

To increase the accuracy of the prediction and therefore further improve the performance of SMVQ, various SMVQ schemes have been developed. Chang and Chen [6] proposed variable-rate side-match finite-state vector quantization with a block classifier. Wei *et al.* [7] reported a three-sided side match technique which uses not only the upper and left sides, but one of the bottom and right sides for the prediction. Yang and Tseng [8] developed a smooth side-match classified vector quantizer which selects state codebook according to the smoothness of the gray levels between neighboring blocks. Chang proposed a gradient match quantization method [9] using gradient match error for the selection of the codewords in the state codebook. The performance of the gradient side match vector quantization method was further improved by combining the non-iterative fractal block coding technique [10,11]. Some other techniques have also been reported for the enhancement of the original SMVQ method. Examples include pattern-based side match VQ [12], side match VQ using gradient based classifier [13], smooth side match weighted method [14] and side match vector quantizers with variable rates according to both the coding quality and the bit rates [15].

In this paper, a neural network based side-match vector quantization method for image coding is presented. In this method, a neural network is designed to predict the variances of the vectors, which are in turn used to select the codewords in the master codebook for the generation of state codebooks for the encoding of the vectors. This technique generates a lower encoding bit rate with a higher reconstructed image quality. Experimental results demonstrate that in terms of PSNR (Peak Signal-to-Noise Ratio) of the reconstructed images, the proposed method significantly outperforms the regular side-match vector quantizer at similar low coding bit-rates.

2. SMVQ FOR IMAGE CODING

As mentioned in previous section, SMVQ takes advantage of both the redundancy within a block and strong correlation between the neighboring blocks for high quality image coding at low bit-rates. The original SMVQ encodes each image block by a smaller-sized state codebook generated from a master codebook using a side-match selection function as shown in Figure 1.

Assume that the master codebook has N codewords with each codeword an $m \times n$ vector denoting by $C_i, i=1, 2, \dots, N$. Also assume that the image to be encoded is partitioned into blocks of size $m \times n$. SMVQ encodes the image blocks in an order from left to right and top to bottom. For each block being encoded, SMVQ uses the side information of its upper and left neighboring blocks to produce the state codebook. The block is encoded as the index of the codeword in the state codebook which is the best match to the block.

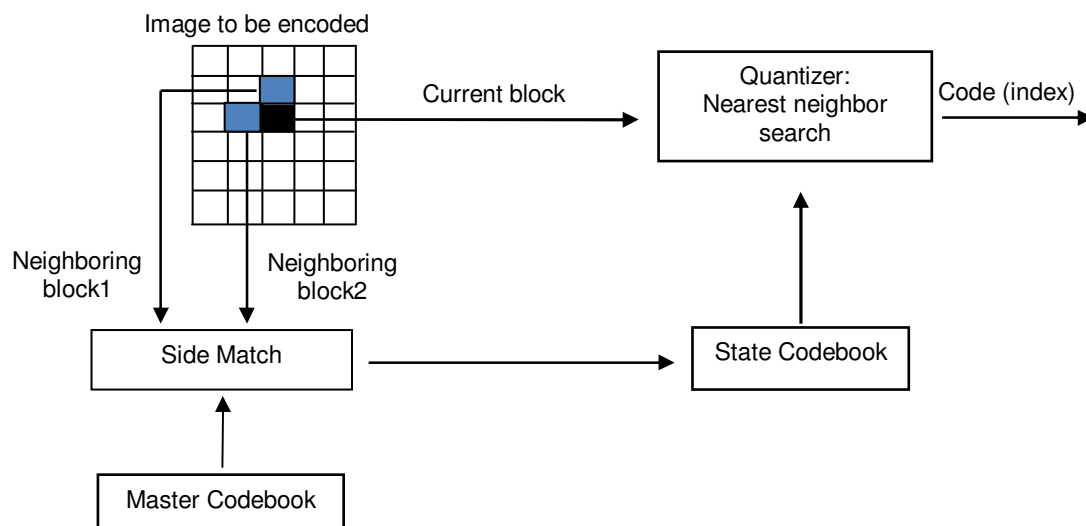


FIGURE 1: Block diagram of SMVQ encoder

Let $x(i, j)$ ($i=1, 2, \dots, m$ and $j=1, 2, \dots, n$) be the pixel in the current image block \mathbf{X} , and $u(i, j)$ ($i=1, 2, \dots, m$ and $j=1, 2, \dots, n$) and $l(i, j)$ ($i=1, 2, \dots, m$ and $j=1, 2, \dots, n$) be its upper block \mathbf{U} and left block \mathbf{L} , respectively. The state codebook is generated and the block is encoded according the following steps:

- (1) For each of the codewords \mathbf{C}_i in the master codebook, calculate the side-match error e_i ,

$$e_i = \sum_{k=1}^n (u(m, k) - c_i(1, k))^2 + \sum_{k=1}^m (l(k, n) - c_i(k, 1))^2 \quad (1)$$

- (2) Select M ($M \leq N$) codewords in the master codebook with the smallest side-match errors as the state codebook codewords \mathbf{S}_i , $i=1, 2, \dots, M$,
- (3) For each of the codewords \mathbf{S}_i in the state codebook, calculate its distortion measure $d_i(\mathbf{S}_i, \mathbf{X})$,

$$d_i(\mathbf{S}_i, \mathbf{X}) = \sum_{k=1}^m \sum_{r=1}^n (s_i(k, r) - x(k, r))^2 \quad (2)$$

- (4) The image block \mathbf{X} is encoded as the index j of the codeword \mathbf{S}_j which produces the smallest distortion $d_j(\mathbf{S}_j, \mathbf{X})$.

For the decoding of each image block, SMVQ first generates the state codebook which was used to encode it according to step (1) and (2) in the encoding process. Once the state codebook is generated, the reconstructed block is obtained by simply selecting the corresponding codeword in the state codebook using the index which is the code of the block generated in the encoding process. The block diagram of SMVQ decoder is shown in Figure 2.

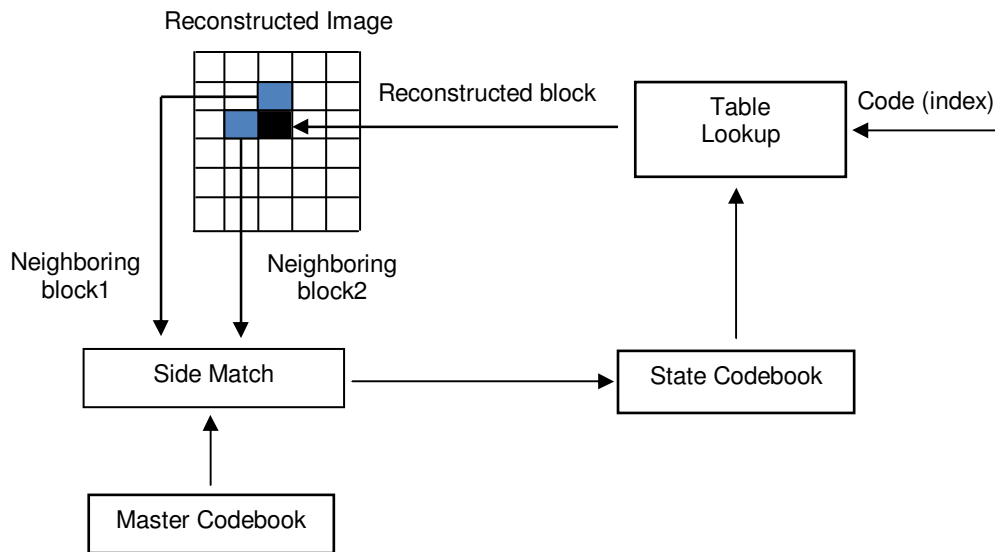


FIGURE 2: Block diagram of SMVQ decoder

3. PROPOSED NEURAL NETWORK BASED SMVQ

3.1 Image Block Variance for SMVQ

SMVQ image coding achieves low bit-rates through using smaller-sized state codebooks generated from master codebook for the coding of the image blocks. The quality of the reconstructed images coded at low bit-rates relies on the accuracy of the prediction of the blocks. An accurate prediction of a block may keep the codeword, which is the best match of the block when a full search in the master codebook is conducted, in the state codebook generated by the side match function. The more accurate prediction makes it possible to select a smaller-sized state codebook for the coding so that the image can be coded at a lower bit-rate with a quality

similar to the one obtained when the master codebook is used. However, the prediction may be inaccurate, especially when the block is within an edge area. An inaccurate prediction on the other hand may keep the codeword which is the best match of the block in a full search out of the generated state codebook when a small size is chosen, and therefore results in a degraded reconstructed image quality.

Since edge and non-edge blocks may have different variances with an edge block highly likely to have a higher variance, the best matched codeword for each block in a full search may scatter around the codeword in the master codebook whose variance is equal to or close to the variance of the block to be encoded, assume that the master codebook is sorted according to the variances of the codewords. Figure 3 shows the distribution of the best matched codewords in full search for the encoding of image Lena using a codebook of size 256 which is sorted by variance of the codewords. The x-axis is the difference value of the index of codeword whose variance is equal or close to the variance of the block being encoded and the index of its best matched codeword in a full search, and the y-axis is the number of image blocks associated with the same index difference value.

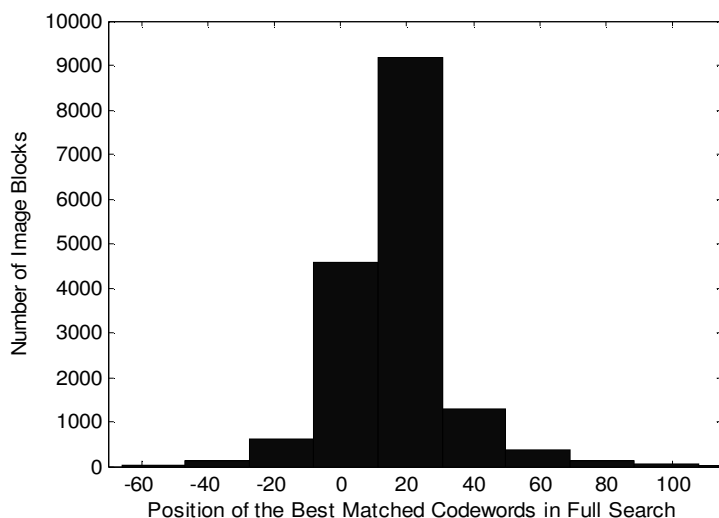


FIGURE 3: Distribution of the best matched codewords in full search

From Figure 3, it can be seen that more than 90% of the best matched codewords are distributed in the range from 6 codewords before to 54 codewords after the codeword whose variance is equal to or close to the variance of the block to be coded. In other words, if the 60 codewords around the codeword whose variance matches the variance of the block being coded are used to be side-matched to generate the state codebook, then there is a much higher possibility that the best matched codeword will be in the state codebook even if the size of the state codebook is selected to be smaller than that selected in the regular SMVQ. This will result in a faster processing speed since a subset of codewords in the master codebook instead of the whole is used for side matching. Furthermore, the coding bit rate can be further lowered due to the fact that smaller size of state codebooks can be used while preserving similar reconstructed image quality.

Based on the above observation, in our proposed SMVQ method, the variances of the image blocks, which are usually sensitive to image edges, are combined into the side matching and state codebook generating process for more accurate prediction and therefore lowered coding bit-rate with preserved image quality.

3.2 Neural Network Based Variance Predictor

In the proposed method, the variance of the block being coded is used to select a subset of the codewords in the master codebook for the side matching and state codebook generation. To implement the process, the variance of the block must be known. In the proposed algorithm, the variance is predicted by a three-layered feed-forward neural network [16] as shown in Figure 4.

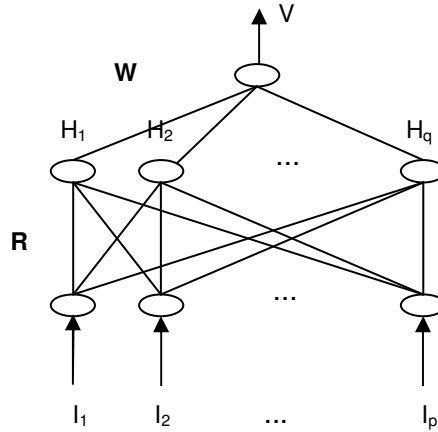


FIGURE 4: Neural network structure of the variance predictor

This neural network contains an input layer with p external inputs, a hidden layer with q neurons and an output layer with one neuron. The external inputs of the network consist of the mean value k of the pixels in the last two rows of the upper block and last two columns of the left block of the current block and the difference values z_i ($i=1, 2, \dots, p-1$) between the mean k and each pixel value in those rows and columns, denoted as $I=(I_1, I_2, \dots, I_p)=(k, z_1, z_2, \dots, z_{p-1})$, where $p = 2*(m+n)+1$, k is calculated as,

$$k = \frac{1}{2*(m+n)} \left(\sum_{i=m-1}^m \sum_{j=1}^n u(i, j) + \sum_{i=1}^m \sum_{j=n-1}^n l(i, j) \right) \tag{3}$$

The output of the network is the predicted variance value of the current block, which is formulated as,

$$V = f \left(\sum_{j=1}^q w_j f \left(\sum_{i=1}^p r_{ij} I_i \right) \right) \tag{4}$$

where V is the output of the network, I_i is the i -th element of the input vector I , r_{ij} is the weight of the connection between the i -th neuron in the input layer and the j -th neuron in the hidden layer, w_j is the connection weight between the j -th neuron in the hidden layer and the neuron in the output layer, and $f(\cdot)$ is the activation function of the neurons, which is

$$f(x) = \frac{2}{1 + e^{-\lambda x}} - 1 \tag{5}$$

where $\lambda > 0$ is the neuron activation function coefficient determining the steepness of the function. The network is trained using back-propagation learning algorithm [16] with sample images. After trained, it can be used to predict the variance of each block being encoded.

3.3 The Encoding and Decoding Algorithms

Given the information presented in the previous sections, the proposed neural network based SMVQ encoder and decoder algorithms are described in this section. The encoder consists of four elements including neural network variance predictor, candidate codeword selector, side

match function, and quantizer as shown in Figure 5. The encoder algorithm can be implemented as follows:

- (1) Sort the master codebook by the variances of the codewords in an ascending order,
- (2) Use the neural network variance predictor to predict the variance of the current block,
- (3) Select T (predefined) ($T < N$) codewords in the master codebook around the codeword whose variance is equal to or is closest to the predicted variance value,
- (4) For each of the codewords C_i selected in step (3), calculate the side-match error e_i using equation (1),
- (5) Select M ($M \leq T$) codewords with the smallest side-match errors from the T codewords selected in step (3) as the state codebook codewords $S_i, i=1, 2, \dots, M$,
- (6) For each of the codewords S_i in the state codebook, calculate its distortion measure $d_i(S_i, X)$ using equation (2),
- (7) If the smallest distortion $d_i(S_i, X)$ is greater than a predefined threshold, the block is coded using full search. Otherwise, it is encoded as the index j of the codeword S_j which produces the smallest distortion $d_j(S_j, X)$,
- (8) Repeat step (2) – (7) for next block encoding.

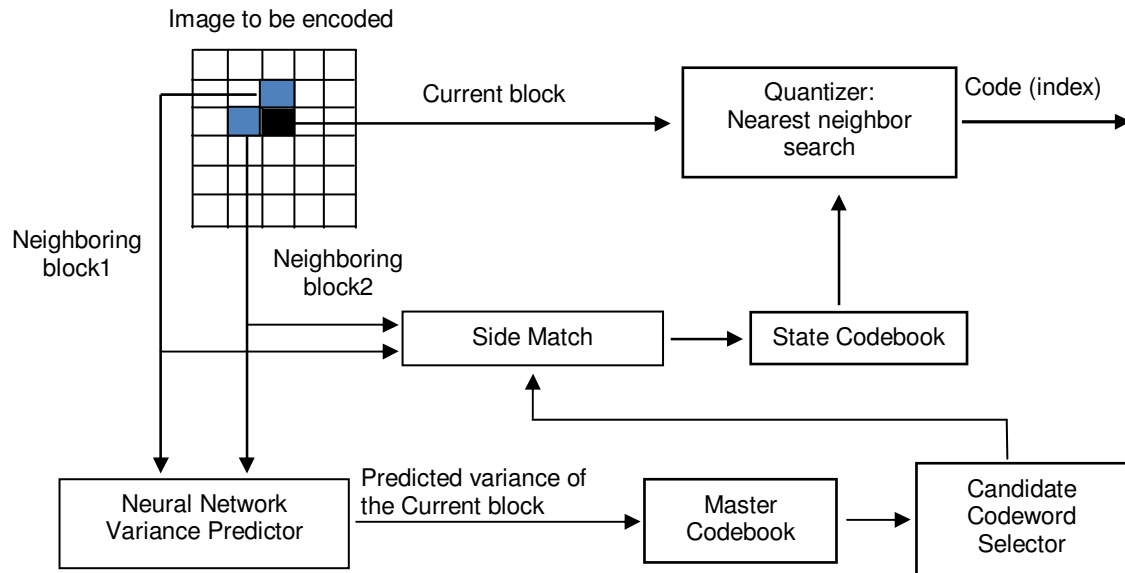


FIGURE 5: Block diagram of the proposed neural network based SMVQ encoder

The structure of the decoder is same as that of the encoder except that the quantizer element in the encoder is replaced by the table lookup element. The block diagram of the decoder is shown as in Figure 6. For the decoding of each image block, the coder first generates the state codebook which was used to encode it according to step (2) and (5) in the encoding process. Once the state codebook is generated, the reconstructed block is obtained by simply selecting the corresponding codeword in the state codebook using the index which is the code of the block generated in the encoding process.

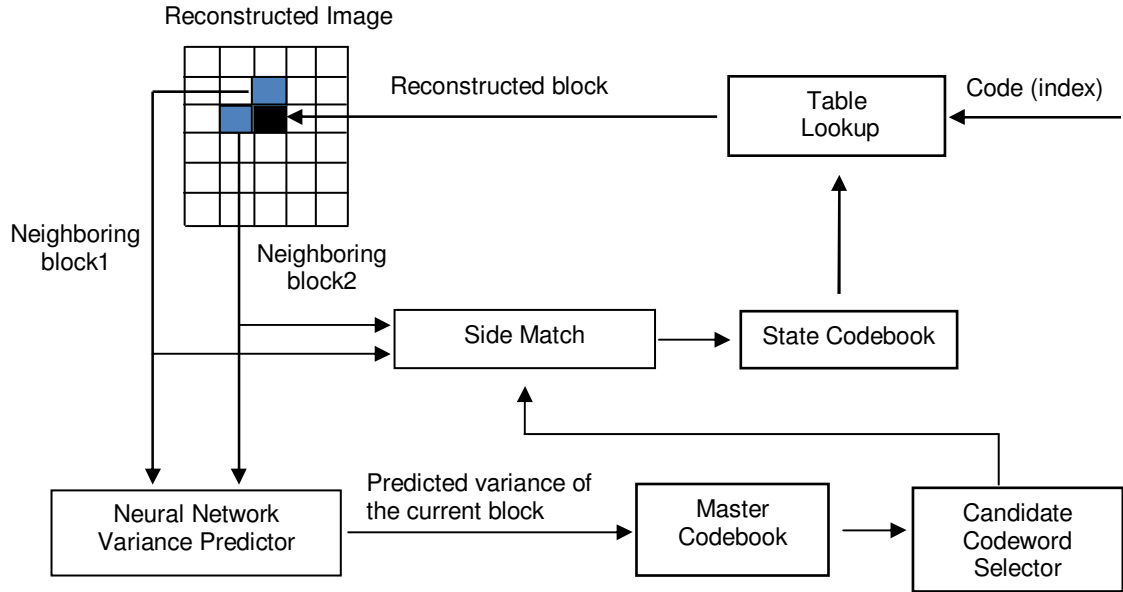


FIGURE 6: Block diagram of the proposed neural network based SMVQ decoder

4. EXPERIMENTS AND DISCUSSION

In this section, the performance of the proposed algorithm is evaluated by experiments. The peak signal-to-noise ratio (PSNR) is used as a quantitative measure of the quality of the reconstructed images. PSNR can be calculated as,

$$PSNR = 10 * \log_{10} \frac{255^2}{\frac{1}{B * B} \sum_{i=1}^B \sum_{j=1}^B (\hat{z}(i, j) - z(i, j))^2} \quad (6)$$

where $\hat{z}(i, j)$ and $z(i, j)$ are the (i, j) th pixels in the reconstructed and original images, respectively, and the images are of size $B \times B$.

In our experiments, the master codebook of size 256 was generated using LBG algorithm [17] and the neural network based variance predictor was trained using 8-bit gray level image Lena of size 512×512 . The trained neural network predictor and coding algorithm were then applied to code the image Lena as well as some other gray images of different objects at various bit rates. For comparison purpose, the corresponding images were also coded by using the regular SMVQ method.

Table 1 shows the PSNRs of the reconstructed images coded around bit rate 0.24bpp (bit per pixel), 0.31bpp and 0.37bpp by using the regular SMVQ and the proposed method. This table indicates that the proposed algorithm improves the PSNR of the reconstructed image over the regular SMVQ average by 1.66dB, 1.06dB and 0.42dB when bit rate is around 0.24bpp, 0.32bpp, and 0.37bpp, respectively. These results demonstrate that in terms of PSNR, the proposed algorithm significantly outperforms the regular SMVQ method, especially when the coding bit rate is low.

	Bit rate: around 0.24bpp		Bit rate: around 0.31bpp		Bit rate: around 0.37bpp	
	SMVQ	Proposed Method	SMVQ	Proposed Method	SMVQ	Proposed Method
Lena	28.70	31.17	30.27	31.83	31.38	31.93
Peppers	27.87	28.86	28.66	29.20	29.06	29.25
Elaine	28.32	29.35	29.14	29.82	29.60	29.86
Airplane	26.50	28.19	27.34	28.43	28.06	28.46
Flowers	30.31	31.75	31.67	32.53	32.38	32.65
Kids	26.15	28.51	27.32	28.95	28.18	29.02
Average	27.98	29.64	29.07	30.13	29.78	30.20

TABLE 1: PSNRs (dB) comparison (Coding Bit Rate around 0.24bpp, 0.31bpp and 0.37bpp)

Figure 7 shows the curves of PSNRs as a function of coding bit-rate for the proposed and regular SMVQ methods. From the figure, it can be observed that from low to high bit-rates, the PSNRs of the images coded by using the proposed method approach much faster than the regular SMVQ method to a constant value, which is the highest PSNR value SMVQ methods can reach and can be obtained when the images are coded using the master codebook. This also confirms that the proposed method has a more accurate prediction of the block being coded and therefore allows to use smaller state codebooks which generate lower coding bit-rates with higher images quality.

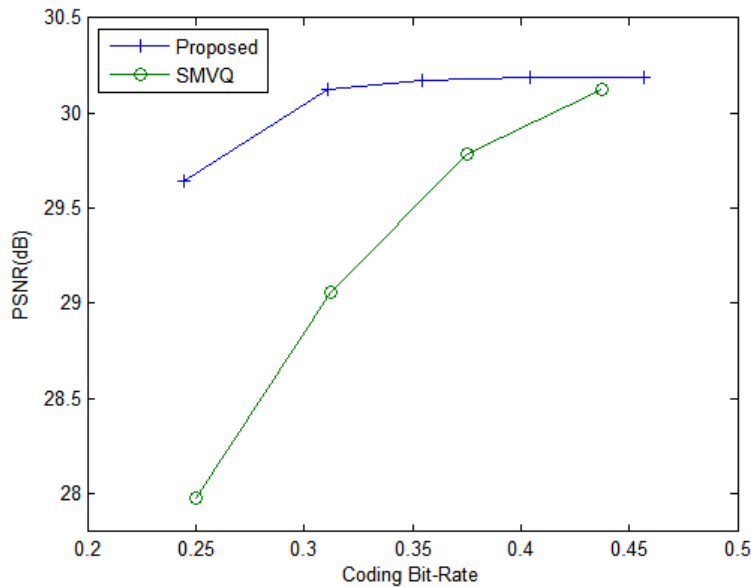


FIGURE 7: PSNR as a function of coding bit-rate

The visual quality of the reconstructed images is also examined. Figure 8 shows the original and corresponding Lena images coded by master codebook at 0.5bpp, the proposed method at 0.30bpp and the regular SMVQ method at 0.31bpp. Magnified face portions of the original and the reconstructed Lena images are shown in Figure 9 for further comparison. These figures show that the images coded by the proposed method are at the quality similar to those coded by the master codebook and they are sharper and visually look much more pleasant than the ones

coded by the regular SMVQ method at the similar coding bit-rate. This is consistent with the PSNR results shown in Table I.



FIGURE 8: (a) Original Lena image; (b) Lena image coded by using the master codebook at 0.5bpp; (c) Lena image coded by using the proposed method at 0.3bpp; (d) Lena image coded by using the regular SMVQ at 0.31bpp

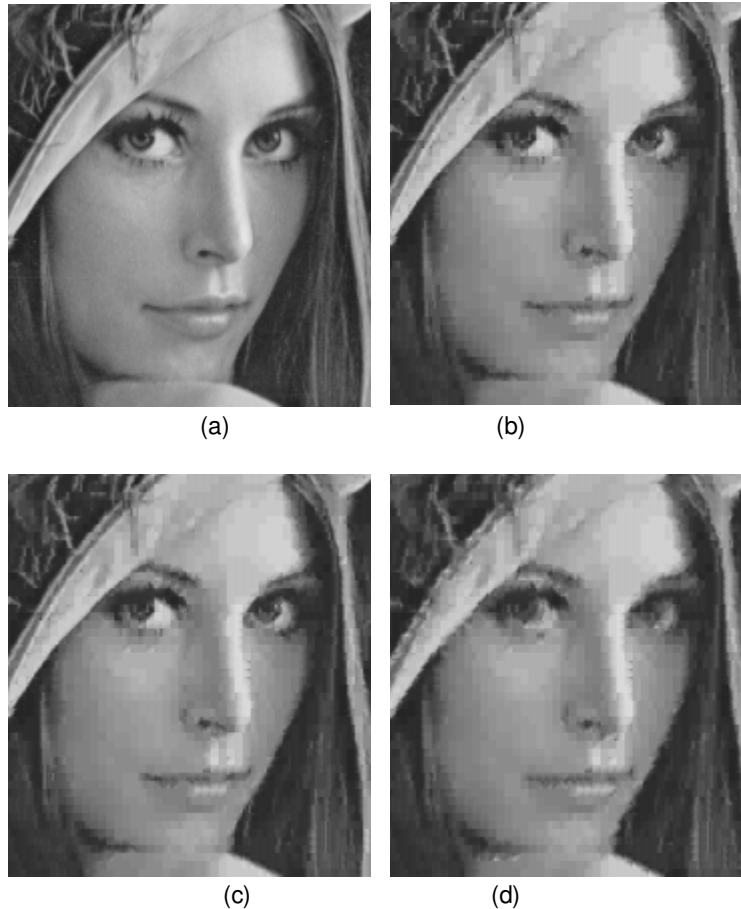


FIGURE 9: Magnified face portion from (a) Original Lena image; (b) Lena image coded by using the master codebook at 0.5bpp; (c) Lena image coded by using the proposed method at 0.3bpp; (d) Lena image coded by using the regular SMVQ at 0.31bpp

5. CONCLUSION

We have presented a new side-match vector quantization method for image coding using a neural network-based variance predictor. In this method, a neural network is used to predict the variances of the image blocks. The predicted variances are in turn used to select a subset of the codewords in the master codebook for the side matching to generate the state codebooks for the encoding of the image blocks. With the involvement of the image block variances, the regular SMVQ is enhanced with more accurate prediction of the current block by using side information and therefore generates smaller state codebooks which encode the images at a lower bit rate with higher image quality. Experimental results demonstrate that in terms of PSNR (Peak Signal-to-Noise Ratio) of the reconstructed images, the proposed method significantly outperforms the regular side-match vector quantizer at similar low coding bit-rates.

6. REFERENCES

- [1] R. M. Gray, "Vector quantization", IEEE ASSP Magazine 1, pp. 4-29, 1984.
- [2] A. Gersho, R. M. Gray, Vector Quantization and Signal, Compression, Kluwer Academic Publishers, 1992.
- [3] N. M. Nasrabadi, R. A. King, "Image coding using vector quantization: a review", IEEE Tran. Communications, vol. 36, no. 8, pp. 957-971, 1988.
- [4] K. Sayood, Introduction to data compression, Morgan Kaufmann Publishers, San Francisco, CA 1996.

- [5] T. Kim, "Side match and overlap match vector quantizers for images", IEEE Trans. Image Process., vol. 1, no. 2, pp. 170 -185, 1992.
- [6] R. F. Chang and W. -T. Chen, "Image coding using variable-rate side-match finite-state vector quantization", IEEE Tran. Image Processing, vol. 2, no. 1, pp. 104-108, 1993.
- [7] H. Wei, P. Tsai and J. Wang, "Three-sided side match vector quantization", IEEE Trans. Circuits and Systems for Video Technology, vol. 10, no. 1, pp. 51–58, 2000.
- [8] S. B. Yang and L. Y. Tseng, "Smooth side-match classified vector quantizer with variable block size", IEEE Tran. Image Processing, vol. 10, no. 5, pp. 677-685, 2001.
- [9] H. T. Chang, "Gradient match vector quantizers for images", Opt. Eng., vol. 39, no. 8, pp.2046-2057, 2000.
- [10] H. T. Chang, "Gradient match and side match fractal vector quantizers for images", IEEE Trans. Image Process., vol. 11, no. 1, pp. 1-9, 2002.
- [11] H. T. Chang and C. J. Kuo, "Iteration-free fractal image coding based on efficient domain pool design", IEEE Tran. Image Processing, vol. 9, pp.329-339, 2000.
- [12] C. C. Chang, F. C. Shiue, T. S. Chen, "Pattern-based side match vector quantization for image compression", Imaging Science Journal, vol. 48, no. 2, pp. 63-76, 2000.
- [13] Z. M. LU, B. Yang, S. H. SUN, "Image compression algorithms based on side-match vector quantizer with gradient-based classifiers", IEICE TRAN. Information and Systems, vol. E85-D, no.9, pp.1409-1415, 2002.
- [14] S. B. Yang, "Smooth side-match weighted vector quantiser with variable block size for image coding", IEE Proc. Vis. Image Signal Processing, vol. 152, no. 6, pp. 763-770, 2005.
- [15] S. B. Yang, "New variable-rate finite state vector quantizer for image coding", Opt. Eng., vol. 44, no. 6, 067001, 2005.
- [16] M. H. Hassoun, Fundamentals of Artificial Neural Network, MIT Press, Cambridge, MA, 1995.
- [17] Y. Linde, A. Buzo and R. M. Gray, "An algorithm for vector quantization design", IEEE Trans. Communications, vol. 28, pp. 84-95, 1980.

A Spectral Domain Dominant Feature Extraction Algorithm for Palm-print Recognition

Hafiz Imtiaz

*Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh*

hafiz.imtiaz@live.com

Shaikh Anowarul Fattah

*Bangladesh University of Engineering and Technology
Dhaka-1000, Bangladesh*

sfattah@princeton.edu

Abstract

In this paper, a spectral feature extraction algorithm is proposed for palm-print recognition, which can efficiently capture the detail spatial variations in a palm-print image. The entire image is segmented into several spatial modules and the task of feature extraction is carried out using two dimensional Fourier transform within those spatial modules. A dominant spectral feature selection algorithm is proposed, which offers an advantage of very low feature dimension and results in a very high within-class compactness and between-class separability of the extracted features. A principal component analysis is performed to further reduce the feature dimension. From our extensive experimentations on different palm-print databases, it is found that the performance of the proposed method in terms of recognition accuracy and computational complexity is superior to that of some of the recent methods.

Keywords: Spectral Feature extraction, Principal Component Analysis (PCA), Two-dimensional Fourier Transform, Classification, Palm-print Recognition, Entropy, Modularization.

1. INTRODUCTION

Conventional ID card and password based identification methods, although very popular, are no more reliable as before because of the use of several advanced techniques of forgery and password-hacking. As an alternative, biometrics, such as palm-print, finger-print, face and iris being used for authentication and criminal identification [1]. The main advantage of biometrics is that these are not prone to theft and loss, and do not rely on the memory of their users. Moreover, they do not change significantly over time and it is difficult for a person to alter own physiological biometric or imitate that of other person's. Among different biometrics, in security applications with a scope of collecting digital identity, the palm-prints are recently getting more attention among researchers [2, 3].

Palm-print recognition is a complicated visual task even for humans. The primary difficulty arises from the fact that different palm-print images of a particular person may vary largely, while those of different persons may not necessarily vary significantly. Moreover, some aspects of palm-prints, such as variations in illumination, position, and scale, make the recognition task more complicated [4].

Palm-print recognition methods are based on extracting unique major and minor line structures that remain stable throughout the lifetime. In this regard, generally, either line-based or texture-based feature extraction algorithms are employed [5]. In the line-based schemes, generally, different edge detection methods are used to extract palm lines (principal lines, wrinkles, ridges, etc.) [6, 7]. The extracted edges, either directly or being represented in other formats, are used for template matching. In cases where more than one person possess similar principal lines, line based algorithms may result in ambiguous identification. In order to overcome this limitation, the texture-based feature extraction schemes can be used, where the variations existing in either the

different blocks of images or the features extracted from those blocks are computed [8, 9]. In this regard, generally, principal component analysis (PCA) or linear discriminant analysis (LDA) are employed directly on palm-print image data or some popular transforms, such as Fourier and discrete cosine transforms (DCT), are used for extracting features from the image data. Given the extracted features, various classifiers, such as decision-based neural networks and Euclidean distance based classifier, are employed for palm-print recognition [6, 7]. Despite many relatively successful attempts to implement face or palm-print recognition system, a single approach, which combines accuracy, robustness, and low computational burden, is yet to be developed.

In this paper, the main objective is to extract precisely spatial variations from each segment of the entire palm-print image instead of considering a global variation pattern. An efficient feature extraction scheme using two dimensional Fourier transform is developed, which operates within those spatial modules to obtain dominant spectral features. It is shown that the discriminating capabilities of the proposed features, that are extracted from the sub-images, are enhanced because of modularization of the palm-print image. Moreover, the improvement of the quality of the extracted features as a result of illumination adjustment has also been analyzed. Apart from considering only the dominant spectral features, further reduction of the feature dimension is obtained by employing the PCA. Finally, recognition task is carried out using a distance based classifiers.

2. BRIEF DESCRIPTION OF THE PROPOSED SCHEME

A typical palm-print recognition system consists of some major steps, namely, input palm-print image collection, pre-processing, feature extraction, classification and template storage or database, as illustrated in Fig. 1. The input palm-print image can be collected generally by using a palm-print scanner. In the process of capturing palm images, distortions including rotation, shift and translation may be present in the palm images, which make it difficult to locate at the correct position. Pre-processing sets up a coordinate system to align palm-print images and to segment a part of palm-print image for feature extraction. For the purpose of classification, an image database is needed to be prepared consisting template palm-images of different persons. The recognition task is based on comparing a test palm-print image with template data. It is obvious that considering images themselves would require extensive computations for the purpose of comparison. Thus, instead of utilizing the raw palm-print images, some characteristic features are extracted for preparing the template. It is to be noted that the recognition accuracy strongly depends upon the quality of the extracted features. Therefore, the main focus of this research is to develop an efficient feature extraction algorithm.

The proposed feature extraction algorithm is based on extracting spatial variations precisely from the spatial modules of the palm-print image instead of utilizing the image as a whole. In view of this, a modularization technique is employed first to segment the entire palm-print into several small segments. It should be noted that variation of illumination of different palm-print images of the same person may affect their similarity. Therefore, prior to feature extraction, an illumination adjustment step is included in the proposed algorithm. After feature extraction, a classifier compares two palm-print features and a database is used to store registered templates and also for verification purposes.

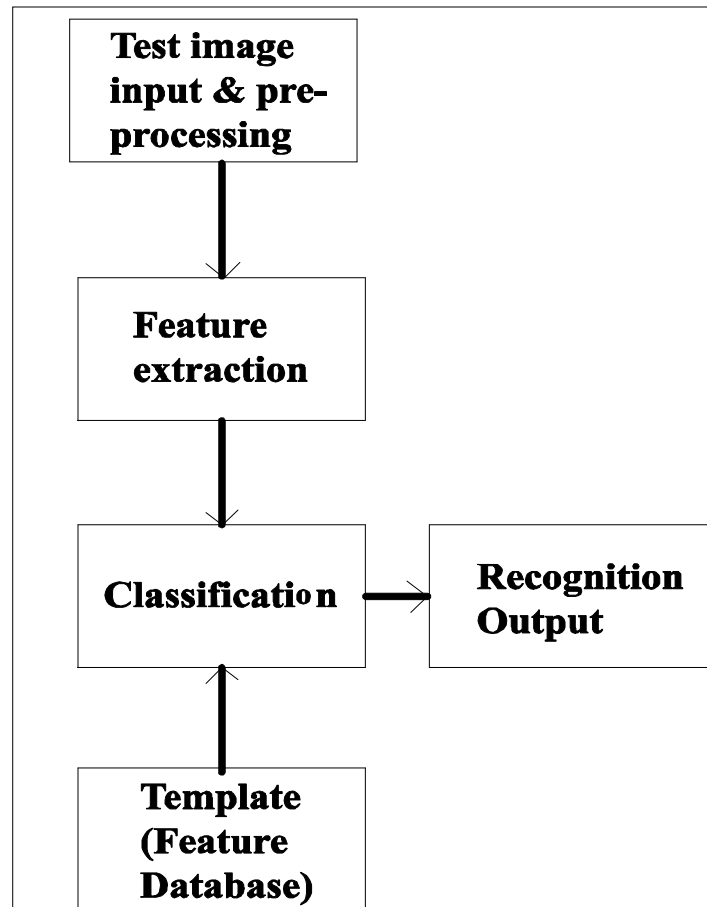


FIGURE 1: Block diagram of the proposed method

3. PROPOSED METHOD

For any type of biometric recognition, the most important task is to extract distinguishing features from the template data, which directly dictates the recognition accuracy. In comparison to person recognition based on face or voice biometrics, palm-print recognition is very challenging even for a human being. For the case of palm-print recognition, obtaining a significant feature space with respect to the spatial variation in a palm-print image is very crucial. Moreover, a direct subjective correspondence between palm-print features in the spatial domain and those in the frequency domain is not very apparent. In what follows, we are going to demonstrate the proposed feature extraction algorithm for palm-print recognition, where spatial domain local variation is extracted from frequency domain transform.

3.1 Spectral Feature Extraction from Spatial Modules

For biometric recognition, feature extraction can be carried out using mainly two approaches, namely, the spatial domain approach and the frequency domain approach. The spatial domain approach utilizes the spatial data directly from the palm-print image or employs some statistical measure of the spatial data. On the other hand, frequency domain approaches employ some kind of transform over the palm-print images for feature extraction and have been widely discussed and applied in image processing [10].

It is well-known that Fourier transform based algorithms offer ease of implementation in practical applications. Hence, we intend to develop an efficient feature extraction scheme using two dimensional Fourier transform. For a function $f(x, y)$ of size $M \times N$ with two-dimensional (2D) variation, the 2D discrete Fourier transform (2D-DFT) is given by [11]

$$\mathcal{F}(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})}, \quad (1)$$

where $u = 1, 2, \dots, M - 1$ and $v = 1, 2, \dots, N - 1$. Palm-prints of a person possess some major and minor line structures along with some ridges and wrinkles. A person can be distinguished from another person based on the differences of these major and minor line structures. Fig. 2 shows sample palm-print images of two different persons. The three major lines of the two persons are quite similar. They differ only in minor line structure. In this case, if we considered the line structures of the two images locally, we may distinguish the two images.

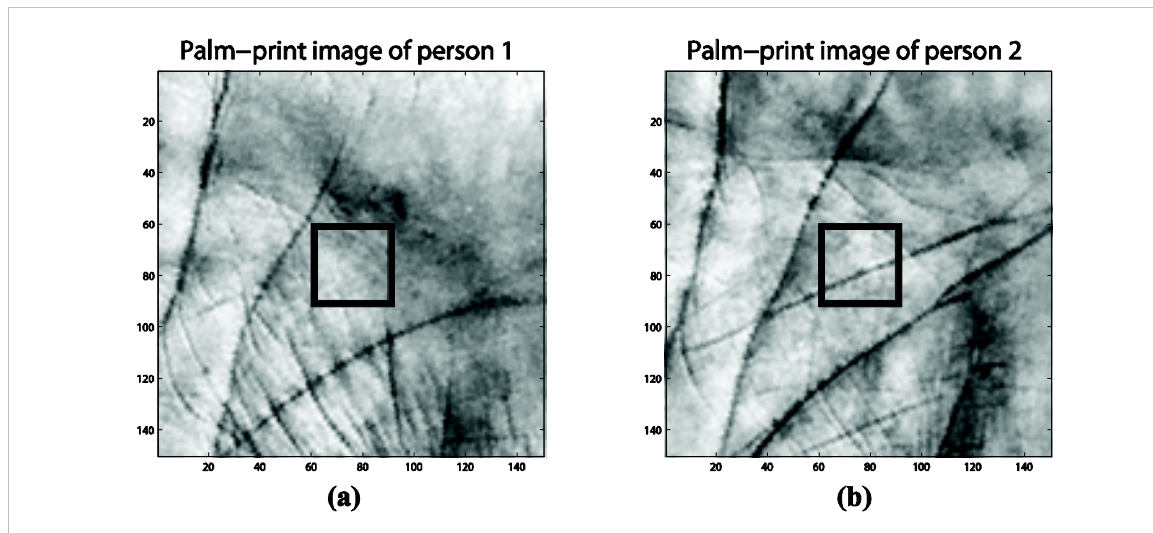


FIGURE 2: Sample palm-print images of two persons. Square block contains portion of images (a) without any minor line (b) with a minor line

For example, if we looked closely within the bordered regions of the palm-print images in Fig. 2, they would seem to be different. Moreover, it is evident that a small block of the palm-print image may or may not contain the major lines but will definitely contain several minor lines. These minor lines may not be properly enhanced or captured when operating on an image as a whole and may not contribute to the feature vector. Hence, in that case, the feature vectors of the palm-prints shown in Fig. 2 may be similar enough to be wrongfully identified as if they belong to the same person. Therefore, we propose to extract features from local zones of the palm-print images.

Figures 3(a) and 3(b) show the 400 lowest 2D-DFT coefficients of the palm-prints images of person 1 and person 2 considered as a whole, respectively. From these figures, it is evident that there exists no significant difference between the two transforms and hence, they are difficult to distinguish, although the palm-print images differ in the bordered region (Fig. 2). On the other hand, Figs. 4(a) and 4(b) show the 400 lowest 2D-DFT coefficients of the bordered regions of the palm-print images of person 1 and person 2, respectively (shown in Fig. 2). In these two figures, the spatial difference in the images is clearly signified in the spectral domain. Next, we compute the Euclidean distance between the DFT coefficients shown in Figs. 3(a) and 3(b) as a measure of similarity. Similarly, the Euclidean distance is computed for the DFT coefficients shown in Figs. 4(a) and 4(b). Fig. 5 shows a comparison between these Euclidean distances. In the former case, where the palm-print image is considered as a whole, the value of the Euclidean distance is smaller than that obtained in the latter case, where only the DFT coefficients of the bordered regions are considered. This clearly indicates that better distinguishable features are extracted from smaller modules than from the entire palm-print image as a whole. It can be observed that within a particular palm-print image, the change in information over the entire image may not be properly captured if the DFT operation is performed upon the image as a whole, because of the

difference in patterns and positions of principal lines, ridges and wrinkles. Even if it is performed, it may offer spectral features with very low between-class separation. In order to obtain high within-class compactness as well as high between-class separability, we propose to segment the palm-print image into some small segments, which are capable of extracting variations in image geometry locally.

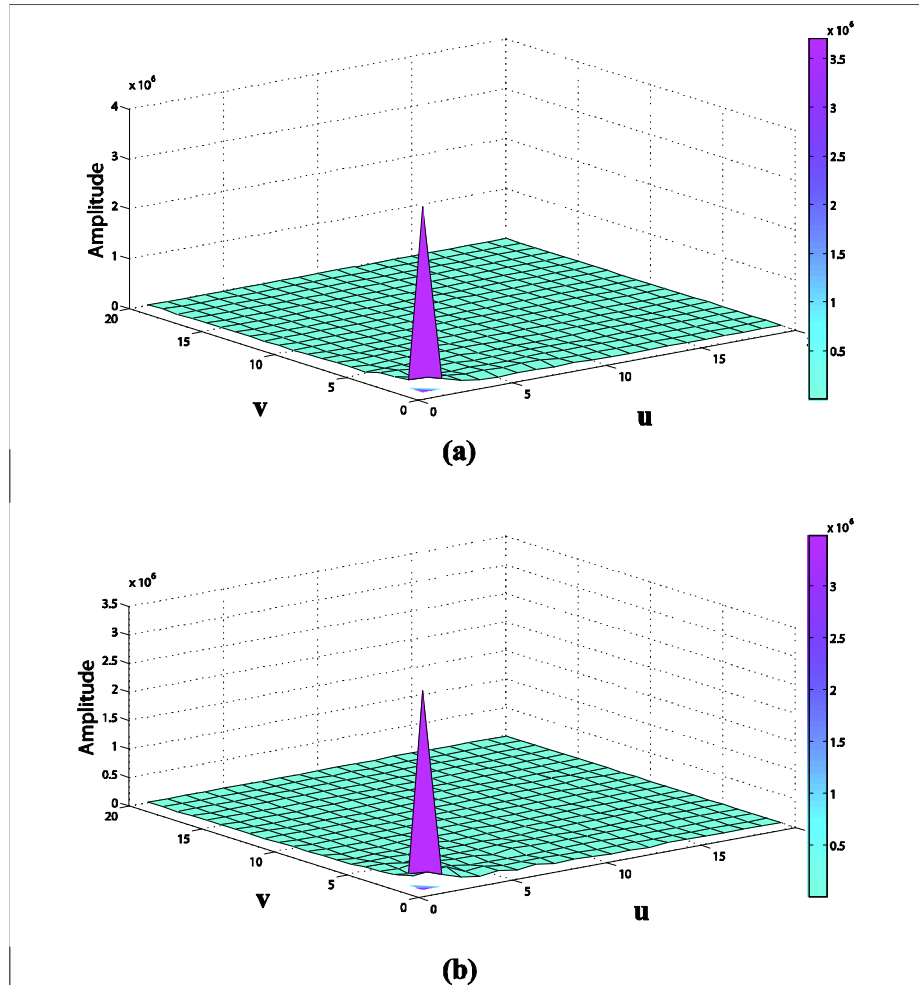


FIGURE 3: 400 lowest 2D-DFT coefficients of the entire palm-prints of (a) Person 1 shown in Fig. 2(a) and (b) Person 2 shown in Fig. 2(b)

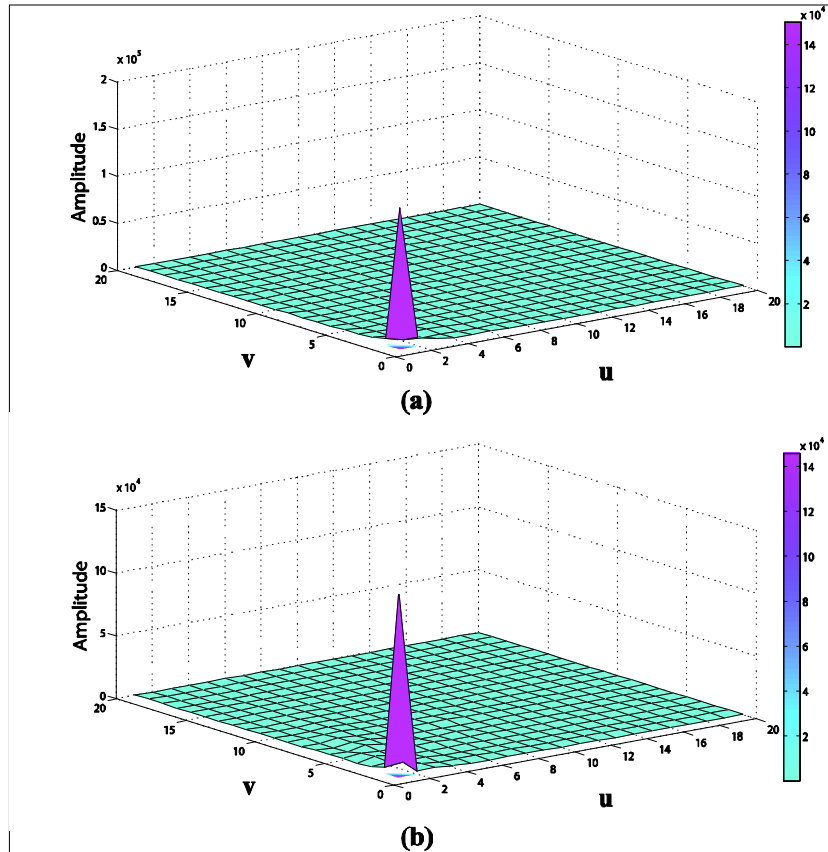


FIGURE 4: 400 lowest 2D-DFT coefficients of the segment of palm-prints of (a) Person 1, corresponding to the square block shown in Fig. 2(a) and (b) Person 2, corresponding to the square block shown in Fig. 2(b)

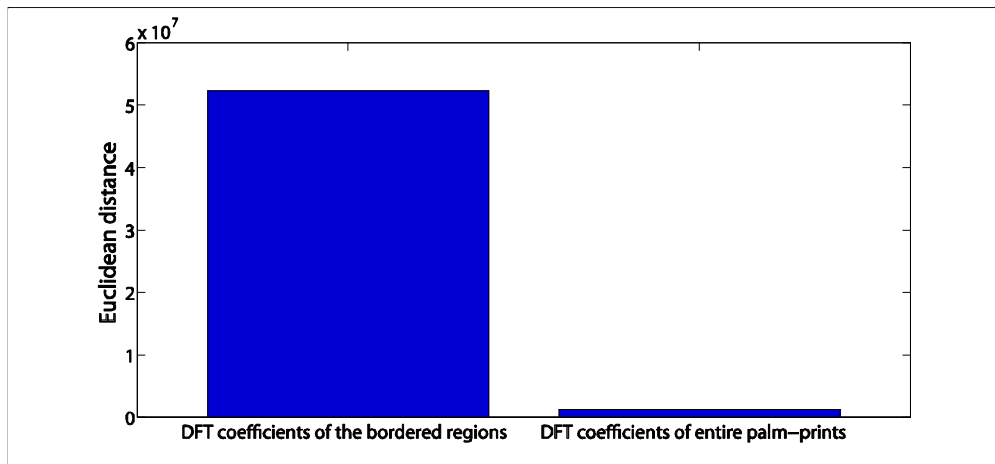


FIGURE 5: Comparison of Euclidean distances of the 2D-DFT coefficients of separate palm-prints shown in Fig. 2

3.2 Effect of Illumination

Light source is one of the key components in the system to capture enough of discriminant information for palm-print recognition. Hence, illumination adjustment is performed to eliminate the dependence of recognition accuracy upon intensity variation. In view of analyzing the effect of

illumination adjustment on the proposed extracted features, in Fig. 6, two sample palm-print images of the same person are shown. The images have different illumination conditions. It is required for a robust feature extraction scheme to ignore the difference in illumination and provide with similar feature vectors for these two images. We perform 2D-DFT upon each image, first without any illumination adjustment and then after performing illumination adjustment. In Fig. 7(a), the correlation of the 2D-DFT coefficients of the two images without any illumination adjustment is shown and in Fig. 7(b), the correlation of the 2D-DFT coefficients of the two images after illumination adjustment is shown. Since the correlation is a measure of similarity, it is evident from these two figures that the latter case exhibits more similarity between the DFT coefficients indicating that the features belong to the same person. In Fig. 8, the similarity measure in terms of Euclidian distance between the 2D-DFT coefficients of the two images for the aforementioned two cases is shown. In case of illumination adjustment, lower value of Euclidean distance is obtained, which clearly indicates a better similarity as well.

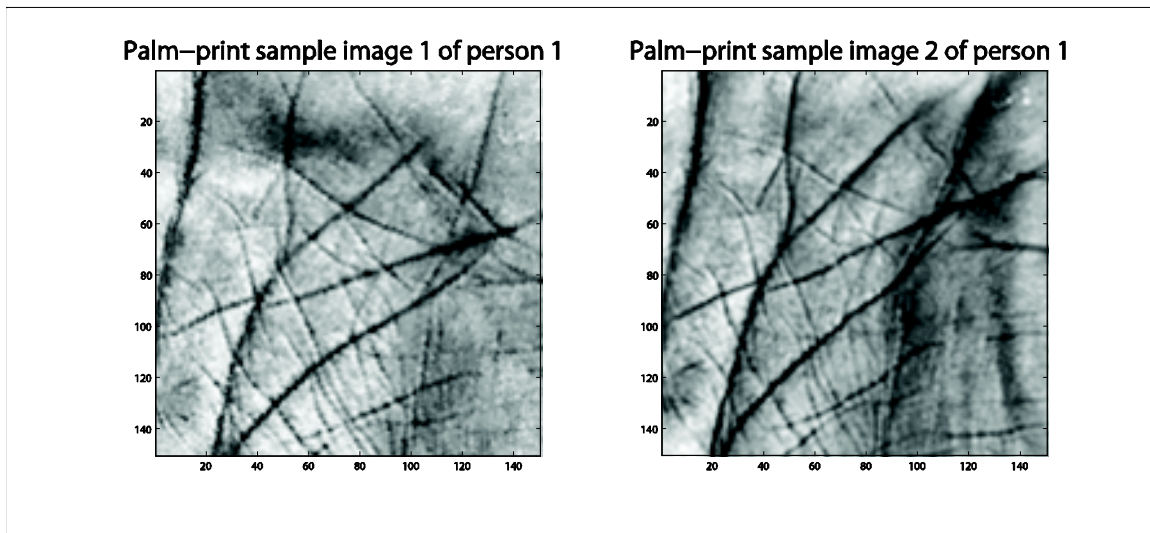


FIGURE 6: Two sample palm-print images of the same person under different illumination

3.3 Proposed Dominant Spectral Feature

In the proposed method, instead of taking the DFT coefficients of the entire image, the coefficients obtained from each module of an image are considered to form a feature vector. However, if all of these coefficients were used, it would definitely result in a feature vector with a very large dimension. In view of reducing the feature dimension, we propose to utilize the magnitudes corresponding to the dominant DFT coefficients as spectral features. The 2D-DFT coefficient corresponding to the maximum magnitude is treated as the dominant coefficient. Considering the magnitudes of the 2D-DFT coefficients in descending order, magnitude values other than the dominant one may also be treated as possible candidates for desired features. In accordance with their magnitude values, these dominant magnitudes are termed as second-dominant (D_2), third-dominant (D_3), and so on. If the magnitude variations along all the segments for the case of different dominant magnitudes remain similar, it would be very difficult to select one of those dominant magnitudes as a desired feature.

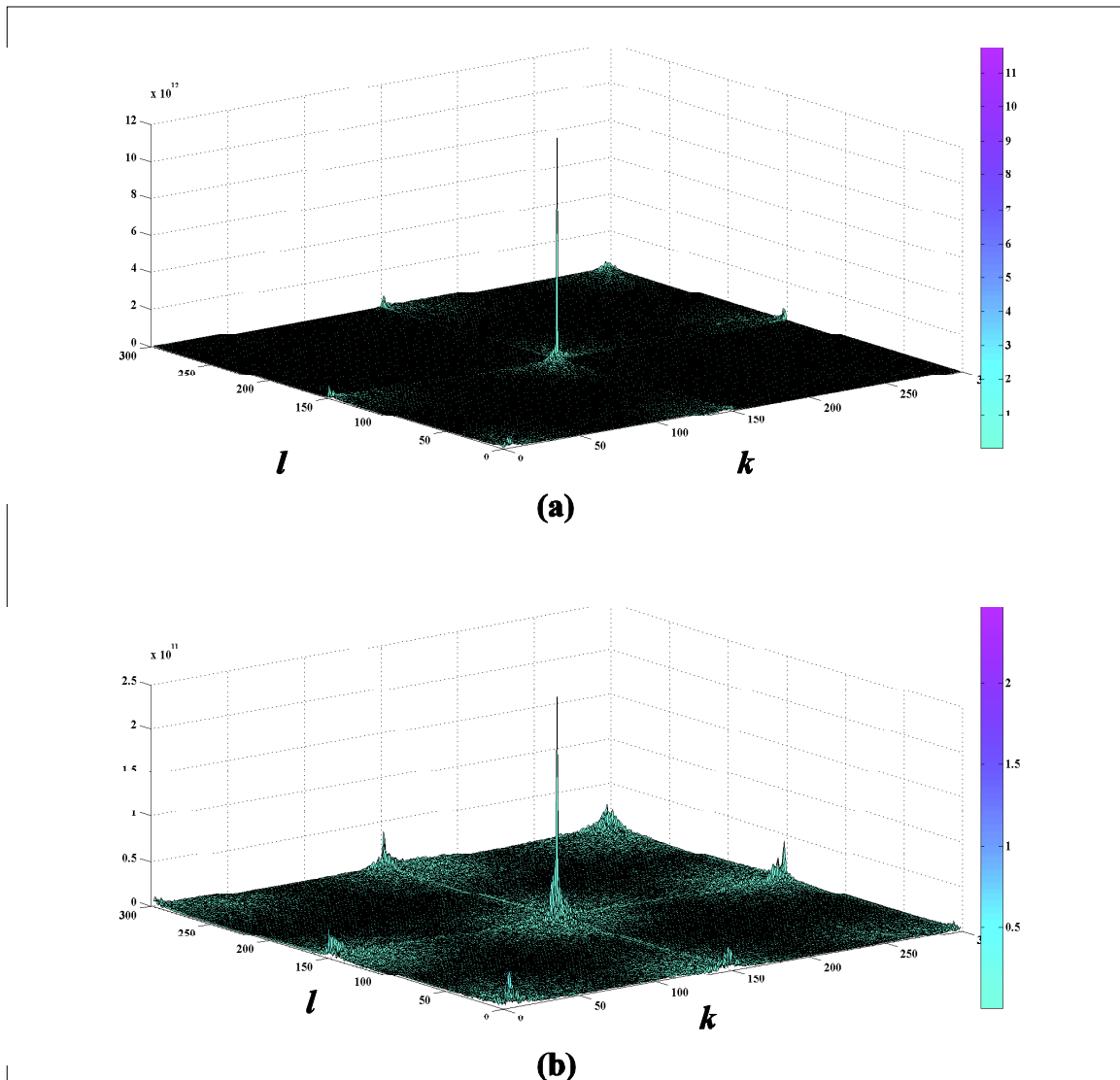


FIGURE 7: Correlation of the 2D-DFT coefficients of the sample palm-print images shown in Fig. 6: (a) with no illumination adjustment and (b) with illumination adjusted

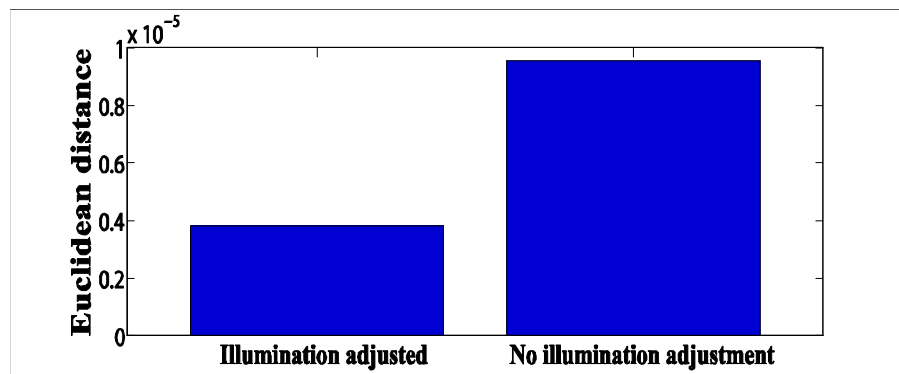


FIGURE 8: Euclidian distance between 2D-DFT coefficients of sample palm-print images shown in Fig. 6

In order to demonstrate the characteristics of the dominant magnitudes in different modules, sample palm-print images of two different persons are shown in Fig. 9. In Fig. 10, four dominant magnitudes ($D1, D2, D3, \text{ and } D4$) obtained from all the modules of the sample palm-print image of Person 1 appeared in Fig. 9 are shown. In this figure, the sample palm-print image is divided into 30 segments. It is found that the first dominant magnitude ($D1$) exhibits completely different characteristics in comparison to other dominant magnitudes. The characteristics of all other dominant magnitudes, in comparison to those of ($D1$), remain almost similar. An analogous behavior is obtained for Person 2 of Fig. 9. It is evident from the figure that $D1$ is the most significant among all the dominant magnitudes and thus, it is sufficient to consider only $D1$ as a desired feature, which also offers an advantage of reduced feature dimension. For a palm-print image of dimension $N \times N$ with M number of segments (with dimension $n \times n$), considering only $D1$ will reduce the length of feature vector from $M \times n \times n$ to M , an order of n^2 reduction. Computing $D1$ in each narrow-width band of a palm-print image, the proposed feature vector is obtained.

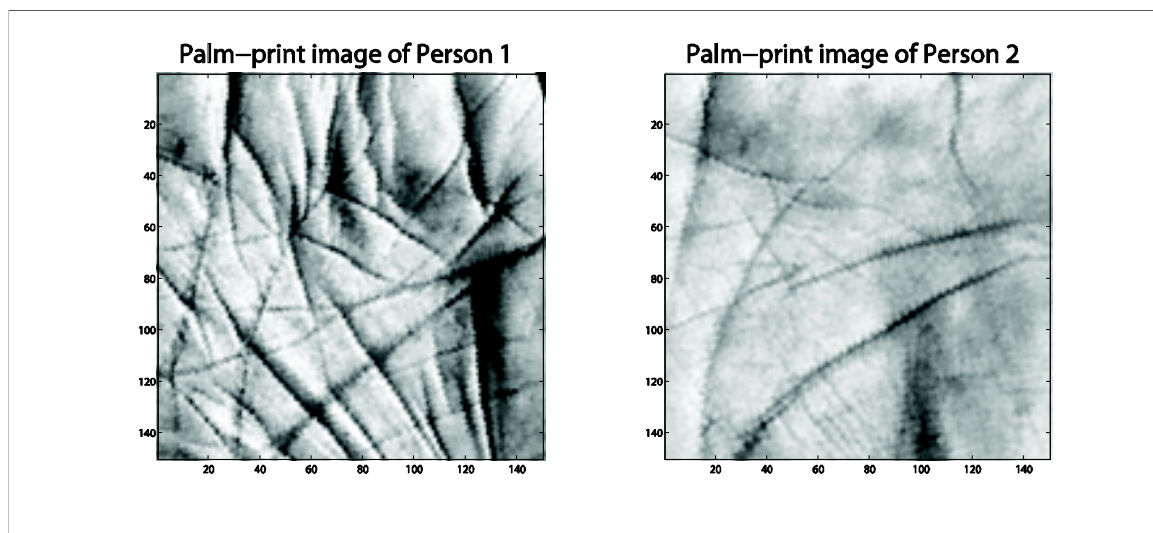


FIGURE 9: Sample palm-print images of two persons

Next, we present an experimental result in order to demonstrate the advantage of extracting the dominant feature ($D1$) from the segments of a palm-print image instead of considering the entire image as a whole. In Fig. 11, centroids of the dominant features obtained from several sample palm-print images of two different persons (as appeared in Fig. 9) are shown considering two different cases: (i) when features are extracted considering the entire palm-print image as a whole and (ii) when features are extracted from each segment separately. It is observed from the figure that, in the first case, the distance between the feature-centroids is extremely small, which strongly discourages to extract a single global variation pattern from the entire image at a time. However, the large between-class-separability in the second case supports the proposed feature selection algorithm, which extracts the features from the entire image considering each local zone separately.

It is observed that a significant variation may occur in the palm-print images of a single person taken under different conditions. In view of demonstrating the effect of such variations on the proposed dominant features, we consider five sample palm-prints for each of the two persons as appeared in Fig. 9. In Fig. 12, the proposed dominant features obtained from different segments of all the sample palm-prints of two different persons are shown. For each person, the centroid of the proposed feature vectors is also shown in the figure (in thick continuous lines). It is to be

noted that the feature centroids of the two different persons are well-separated even though the major lines of the two palm-print images are quite similar considering the pattern and position. It is also observed that a low degree of scattering exists among the features around their corresponding centroids. Hence, the dominant features extracted locally within a palm-print image offer not only a high degree of between-class separability but also satisfactory within-class compactness.

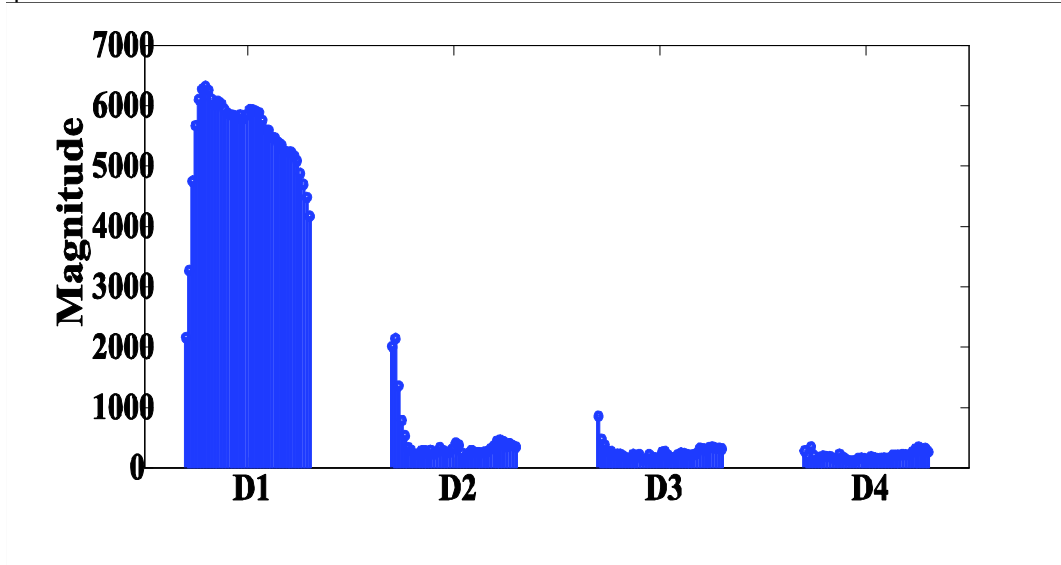


FIGURE 10: Proposed dominant magnitude-features corresponding to palm-prints shown in Fig. 9

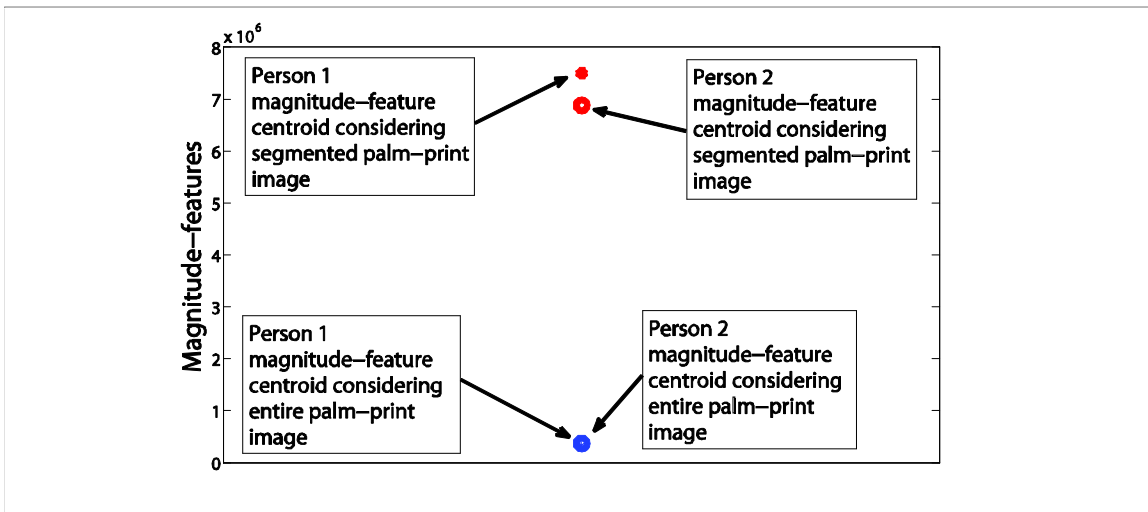


FIGURE 11: Centroids of dominant features corresponding to palm-prints shown in Fig. 9

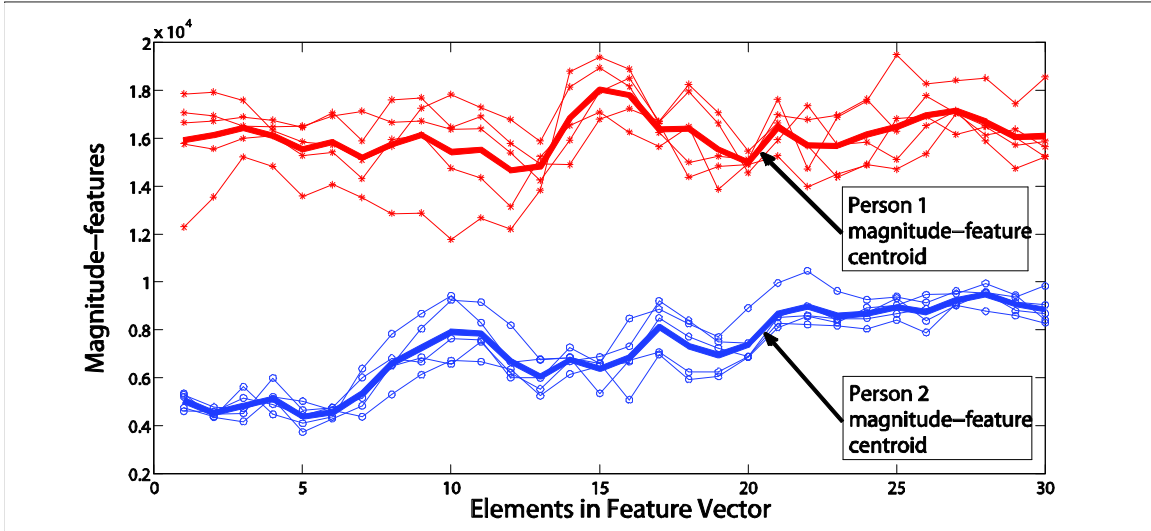


FIGURE 12: Variation of dominant features with segments for several palm-print images of two persons corresponding Fig. 9

3.4 Feature Dimensionality Reduction

For the cases where the acquired palm-print are of very high resolution, even after selection of dominant features from the small segments of the palm-print image, the feature vector length may still be very high. Further dimensionality reduction may be employed for reduction in computational burden.

Principal component analysis (PCA) is a very well-known and efficient orthogonal linear transformation [12]. It reduces the dimension of the feature space and the correlation among the feature vectors by projecting the original feature space into a smaller subspace through a transformation. The PCA transforms the original p -dimensional feature vector into the L -dimensional linear subspace that is spanned by the leading eigenvectors of the covariance matrix of feature vector in each cluster ($L < p$). PCA is theoretically the optimum transform for given data in the least square sense. For a data matrix, X^T , with zero empirical mean, where each row represents a different repetition of the experiment, and each column gives the results from a particular probe, the PCA transformation is given by:

$$Y^T = X^T W = V \Sigma^T \tag{2}$$

where the matrix Σ is an $m \times n$ diagonal matrix with nonnegative real numbers on the diagonal and $W \Sigma V^T$ is the singular value decomposition of X . If q sample palm-print images of each person are considered and a total of M dominant DFT coefficients are selected per image, the feature space per person would have a dimension of $q \times M$. For the proposed dominant spectral features, implementation of PCA on the derived feature space could efficiently reduce the feature dimension without losing much information. Hence, PCA is employed to reduce the dimension of the proposed feature space.

3.5 Distance Based Classifier

In the proposed method, for the purpose of recognition using the extracted dominant features, a distance-based similarity measure is utilized. The recognition task is carried out based on the distances of the feature vectors of the training palm-images from the feature vector of the test palm-image. Given the m -dimensional feature vector for the k -th sample image of the j -th person be $\{y_{jk}(1), y_{jk}(2), \dots, y_{jk}(m)\}$ and a test sample image f with a feature vector

$\{v_f(1), v_f(2), \dots, v_f(m)\}$, a similarity measure between the test image f of the unknown person and the sample images of the j -th person, namely *average sum-squares distance*, D , is defined as

$$D_j^f = \frac{1}{q} \sum_{k=1}^q \sum_{i=1}^m |v_{jk}(i) - v_f(i)|^2, \tag{3}$$

where a particular class represents a person with q number of sample palm-print images. Therefore, according to (3), given the test sample image f , the unknown person is classified as the person j among the p number of classes when

$$D_j^f \leq D_g^f, \forall j \neq g \text{ and } \forall g \in \{1, 2, \dots, p\} \tag{4}$$

4. EXPERIMENTAL RESULTS

Extensive simulations are carried out in order to demonstrate the effectiveness of the proposed method of palm-print recognition using the palm-print images of several well-known databases. Different analyses showing the effectiveness of the proposed feature extraction algorithm have been shown. The performance of the proposed method in terms of recognition accuracy is obtained and compared with those of some recent methods [9, 13].

4.1 Palm-print Databases

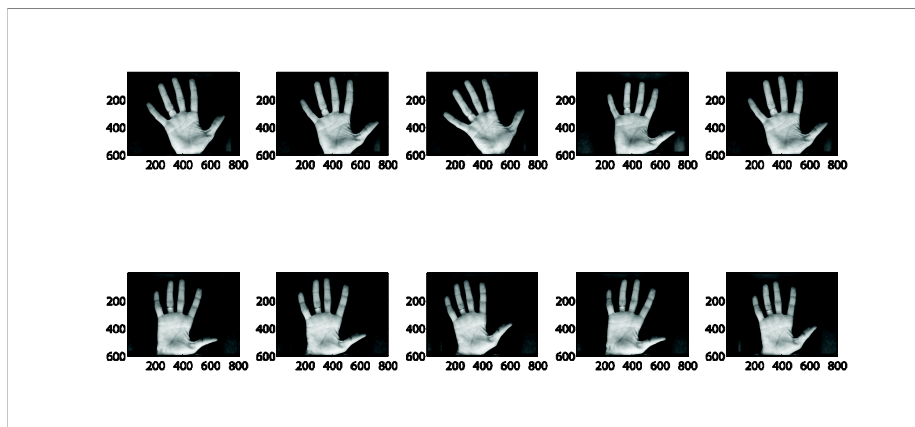


FIGURE 13: Sample palm-print images from the IITD database

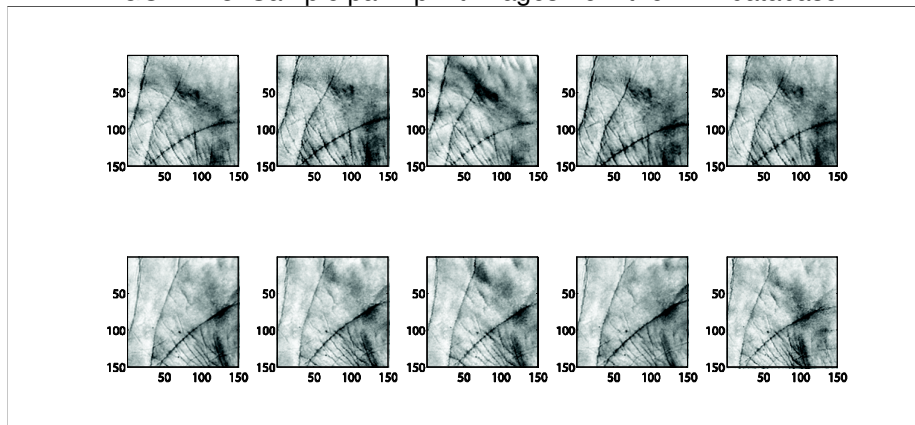


FIGURE 14: Sample palm-print images from the IITD database after cropping

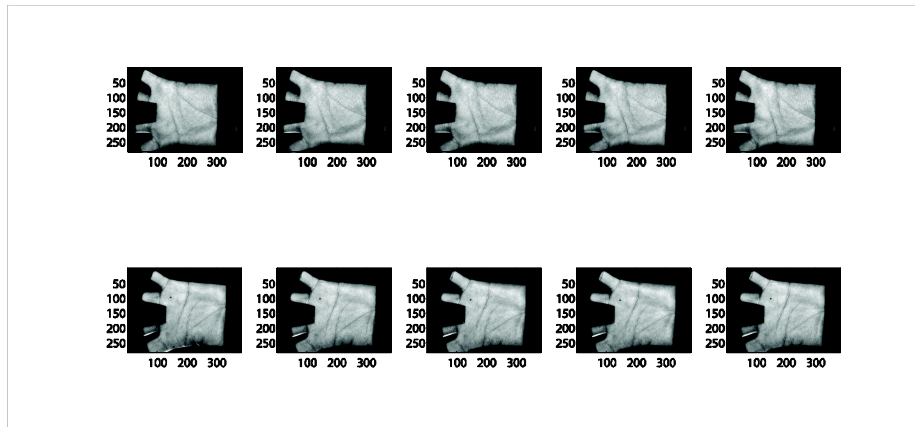


FIGURE 15: Sample palm-print images from the PolyU database

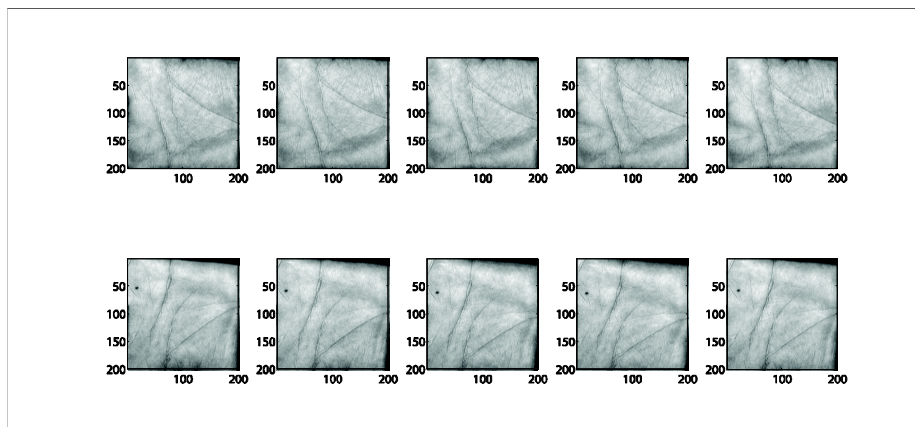


FIGURE 16: Sample palm-print images from the PolyU database after cropping

In this section, palm-print recognition performance obtained by different methods has been presented using two standard databases, namely, the PolyU palm-print database (version 2) [14] and the IITD palm-print database [15]. In Figs. 13 and 15, sample palm-print images from the PolyU database and the IITD database are shown, respectively. The PolyU database (version 2) contains a total of 7752 palm-print images of 386 persons. Each person has 18 to 20 different sample palm-print images taken in two different instances. The IITD database, on the other hand, consists a total of 2791 images of 235 persons, each person having 5 to 6 different sample palm-print images for both left hand and right hand. It can be observed from Figs. 13 and 15 that not all the portions of the palm-print images are required to be considered for feature extraction [2]. The portions of the images containing fingers and the black regions are discarded from the original images to form the regions of interest (ROI) as shown in Figs. 14 and 16.

4.2 Performance Comparison

In the proposed method, dominant spectral features that are obtained from all the modules of a particular palm-print image are used to form the feature vector of that image and the recognition task is carried out on the reduced feature space as described in Section 3.5. The experiments were performed following the leave-one-out cross validation rule. For simulation purposes, the module size for the PolyU database and the IITD database has been chosen as 20x20 pixels and 15x15 pixels, respectively. For the purpose of comparison, recognition accuracy obtained using the proposed method along with those reported in [9] and [13] are listed in Table 1. It is evident from the table that the recognition accuracy of the proposed method is comparatively higher than

those obtained by the other methods. The performance of the proposed method is also very satisfactory for the IITD database (for both left hand and right hand palm-print images). An overall recognition accuracy of 99.91% is achieved.

Method	Recognition Accuracy
Proposed method	99.94%
Method [9]	97.50%
Method [13]	98.00%

TABLE 1: Comparison of recognition accuracies.

5. CONCLUSIONS

In the proposed palm-print recognition scheme, instead of operating on the entire palm-print image at a time, dominant spectral features are extracted separately from each of the modules obtained by image-segmentation. It has been shown that because of modularization of the palm-print image, the proposed dominant spectral features, that are extracted from the sub-images, attain better discriminating capabilities. The proposed feature extraction scheme is shown to offer two-fold advantages. First, it can precisely capture local variations that exist in the major and minor lines of palm-print images, which plays an important role in discriminating different persons. Second, it utilizes a very low dimensional feature space for the recognition task, which ensures lower computational burden. For the task of classification, a Euclidean distance based classifier has been employed and it is found that, because of the quality of the extracted features, such a simple classifier can provide a very satisfactory recognition performance and there is no need to employ any complicated classifier. From our extensive simulations on different standard palm-print databases, it has been observed that the proposed method, in comparison to some of the recent methods, provides excellent recognition performance.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude towards the authorities of the Department of Electrical and Electronic Engineering and Bangladesh University of Engineering and Technology (BUET) for providing constant support throughout this research work.

6. REFERENCES

- [1] A. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4 – 20, 2004.
- [2] A. Kong, D. Zhang, and M. Kamel, "A survey of palmprint recognition," *Pattern Recognition*, vol. 42, pp. 1408–1418, 2009.
- [3] A. Kong, D. Zhang, and G. Lu, "A study of identical twins palmprint for personal verification," *Pattern Recognition*, vol. 39, pp. 2149–2156, 2006.
- [4] C. Han, H. Cheng, C. Lin, and K. Fan, "Personal authentication using palm-print features," *Pattern Recognition*, vol. 36, pp. 371–381, 2003.
- [5] X. Wu, D. Zhang, and K. Wang, "Palm line extraction and matching for personal authentication," *IEEE Trans. Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 36, no. 5, pp. 978 –987, 2006.
- [6] X. Wu, K. Wang, and D. Zhang, "Fuzzy direction element energy feature (FDEEF) based palmprint identification," in *Proc. Int. Conf. Pattern Recognition*, vol. 1, 2002, pp. 95–98.

- [7] S. Kung, S. Lin, and M. Fang, "A neural network approach to face/palm recognition," in Proc. IEEE Workshop Neural Networks for Signal Processing, 1995, pp. 323–332.
- [8] T. Connie, A. Jin, M. Ong, and D. Ling, "An automated palmprint recognition system," Image and Vision Computing, vol. 23, pp. 501–515, 2005.
- [9] M. P. Dale, M. A. Joshi, and N. Gilda, "Texture based palmprint identification using DCT features," in Proc. Int. Conf. Advances in Pattern Recognition, vol. 7, 2009, pp. 221–224.
- [10] W. Li, D. Zhang, L. Zhang, G. Lu, and J. Yan, "3-D palmprint recognition with joint line and orientation features," IEEE Trans. Systems, Man, and Cybernetics, Part C, vol. 41, no. 2, pp. 274 –279, 2011.
- [11] R. C. Gonzalez and R. E. Woods, Digital Image Processing. plus 0.5em minus 0.4em Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1992.
- [12] I. Jolloffe, "Principal component analysis," Springer-Verlag, Berlin, 1986.
- [13] J. Lu, Y. Zhao, and J. Hu, "Enhanced gabor-based region covariance matrices for palmprint recognition," Electron. Lett., vol. 45, pp. 880–881, 2009.
- [14] "The Hong Kong Polytechnic University (PolyU) 2D 3D Palmprint Database." [Online]. Available: <http://www4.comp.polyu.edu.hk/~biometrics/>
- [15] "IIT Delhi Touchless Palmprint Database." [Online]. Available: http://web.iitd.ac.in/~ajaykr/Database_Palm.htm

Quaternion Based Omnidirectional Machine Condition Monitoring System

Wai-Kit Wong

*Faculty of Engineering and Technology,
Multimedia University,
75450 JLN Ayer Keroh Lama,
Melaka, Malaysia.*

wkwong@mmu.edu.my

Chu-Kiong Loo

*Faculty of Computer Science and Information Technology
University of Malaya
50603 Lembah Pantai
Kuala Lumpur, Malaysia.*

ckloommu@gmail.com

Way-Soong Lim

*Faculty of Engineering and Technology,
Multimedia University,
75450 JLN Ayer Keroh Lama,
Melaka, Malaysia.*

wslim@mmu.edu.my

Abstract

Thermal monitoring is useful for revealing some serious electrical problems in a factory that often go undetected until a serious breakdown occurs. In factories, there are various types of functioning machines to be monitored. When there is any malfunctioning of a machine, extra heat will be generated which can be picked up by thermal camera for image processing and identification purpose. In this paper, a new and effective omnidirectional machine condition monitoring system applying log-polar mapper, quaternion based thermal image correlator and max-product fuzzy neural network classifier is proposed for monitoring machine condition in an omnidirectional view. With this monitoring system, it is convenient to detect and monitor the conditions of (overheat or not) of more than one machines in an omnidirectional view captured by using a single thermal camera. Log-polar mapping technique is used to unwarped omnidirectional thermal image into panoramic form. Two classification characteristics namely: peak to sidelobe ratio (PSR) and real to complex ratio of the discrete quaternion correlation output (p-value) are applied in the proposed machine condition monitoring system. Large PSR and p-value observe in a good match among correlation of the input thermal image with a particular reference image, while small PSR and p-value observe in a bad/not match among correlation of the input thermal image with a particular reference image. Simulation results also show that the proposed system is an efficient omnidirectional machine monitoring system with accuracy more than 97%

Keywords: Machine Condition Monitoring System, Neuro Fuzzy System, Thermal Imaging, Quaternion, Omnidirectional.

1. INTRODUCTION

Many factories in all over the world rely on machines to help improve their production and process. An effective machine condition monitoring system play an important role in those factories to ensure that their production and process are running smoothly all the time. Infrared cameras or thermal cameras are used in many heavy factories for monitoring the temperature conditions of the machines. When there is any malfunctioning of machines, extra heat will be generated and it can be picked up by thermal camera. Thermal camera will generate an image to indicate the condition of the machine. This enables the operator to decide on the on/off switch. Any malfunctioned machines detected will proceed to further repairmen action. This process is term as thermal imaging monitoring.

Thermal imaging monitoring is more convenient in compare to conventional maintenance method. In conventional maintenance method, functioning machines needs to be frequently monitored by operator. The problem is it required more man power and longer maintenance time. However, with thermal imaging monitoring, the operating machines can be maintained and monitored by the operator with observing the thermal images captured routinely on the functioning machines and display on a monitor, even from a remote location. So, hands on workload, man power, maintenance time can be reduced and improve safety, since some overheat devices cannot see through eyes, but can read from thermal images, hence the use of thermal imaging monitoring can prevent accident happen too [1].

If a single thermal camera is applied to monitor a single machine, then for more functioning machines in different angle of view, more thermal cameras are required. This will increase cost, beside complicated the monitoring network. In this paper, a new and effective omnidirectional machine condition monitoring system applying log-polar mapper, quaternion based thermal image correlator and max-product fuzzy neural network classifier is proposed for monitoring machine condition in an omnidirectional view. In terms of hardware part, an omnidirectional thermal imaging system consists of thermal camera, custom made IR reflected hyperbolic mirror, camera mirror holder, and laptop/PC as image processing tools is proposed for effective machine condition monitoring purpose. Thermal images captured from the custom made hyperbolic mirror are in omnidirectional view. Log-polar mapping technique [2] is applied for unwarping the captured omnidirectional thermal images into panoramic form, providing the observer or image processing tools a complete wide angle of view.

Quaternion correlator is so far commonly used in color human face recognition [3] and color alphanumeric words recognition [4]. It is found to be useful in machine condition monitoring too, especially in thermal condition monitoring. In [5], quaternion correlator was proposed to be used in thermal image recognition for machine condition monitoring system, so called the quaternion based thermal image correlator. The quaternion based thermal image correlator was in pair with max-product fuzzy neural network classifier were used to monitor fixed angle machines' condition. The experimental results in [5] also shown that the proposed system achieves high accuracy in monitoring machines condition. It's never been used in omnidirectional approach. Therefore, in this paper, quaternion based thermal image correlator and max-product fuzzy neural network classifier are proposed for omnidirectional machine condition monitoring system as a new approach.

In quaternion based thermal image correlator proposed in [5], a strong and sharp peak can be observed in the output correlation plane when the input thermal image comes from the authentic class (input thermal image matches with a particular training/reference image stored in the database), and there will be no discernible peak if the input thermal image comes from imposter class (input thermal image does not match with the particular reference image). For better recognition, peak-to-sidelobe ratio (PSR) [6] is introduced to test whether an input thermal image belongs to the authentic class or not. In [5], it is discovered that by considering the peak value with the region around the peak value is more accurate compare to just a single peak point. According to the results in [5], the higher is the value of PSR, the more likely is the input thermal image belonging to the reference image class. Another parameter use in [5] for quaternion correlation is the real to complex ratio of the discrete quaternion correlation output (p -value [4]). p -value is used in quaternion based thermal image correlator for measuring the quaternion correlation output between the colors, shape, size and brightness of the input thermal image and a particular reference thermal image.

A max-product fuzzy neural network classifier is also proposed to perform classification on the thermal images based on the PSR and p -value output from the quaternion based thermal image correlator. Classification in pattern recognition [7] refers to a procedure whereby individual patterns are placed into groups based on quantitative information on one or more characteristics inherent in the patterns and based on a training set of previously labeled patterns, known as classes. The purpose of classification is to establish a rule whereby a new observed pattern will map into one of the existing classes. These classes are predefined by a partition of the sample

space, which is the attributes themselves. For example, a machine may be classified as overheated if the colors display in the thermal image is brighter than the predetermined limits.

In [5], the max-product fuzzy neural network classifier was modified according to [8] to perform the classification in machine condition monitoring system. According to the designed max-product fuzzy neural network classifier, both the PSR and p -value output from the quaternion based thermal image correlator are first fuzzified with Gaussian membership function. The max-product fuzzy neural network classifier is applied for accurate classification with the weights obtained from training reference images. The weights are then applied for classification of input images in real time application. The same classifier is applied in this paper, for the omnidirectional approach. Experimental results show that in the authentic case, if an input image is well matched with a particular reference image in the database, followed by performing quaternion correlation on these two images, their output correlation plane will have sharp peaks. However in imposter case, if an input thermal image is not matched with a particular reference image in the database, the output correlation plane is flat. Large peak to sidelobe ratio (PSR) and real to complex ratio of the discrete quaternion correlation output (p -value) is proven to have a good match among correlation of the input thermal image with a particular reference image, while small PSR and p -value reflect reversely. Experimental results in this paper also show that the proposed system is an efficient wide angle coverage machine condition monitoring system with accuracy above 97%.

This paper is organized in the following order: Section 2 briefly comments on the quaternion based omnidirectional machine condition monitoring system, section 3 summarize the log-polar image geometry and the mapping techniques. The algorithm of the proposed quaternion based thermal image correlator is described in section 4. Section 5 describes the structure of the max-product fuzzy neural network classifier. In section 6, the experimental results is discussed. Finally section 7 summarizes the work and some suggestions are proposed for future work.

2. QUATERNION BASED OMNIDIRECTIONAL MACHINE CONDITION MONITORING SYSTEM MODEL

The quaternion based omnidirectional machine condition monitoring system developed in this paper is shown in Fig. 1.

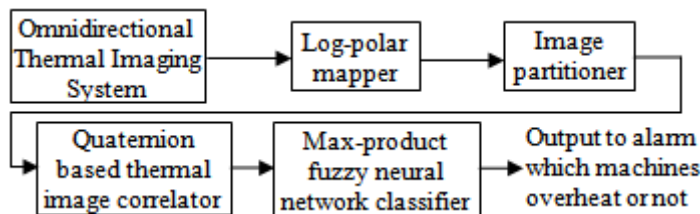


FIGURE 1: Quaternion based omnidirectional machine condition monitoring system

The omnidirectional thermal imaging system consists of three elements, which are the custom made IR reflected hyperbolic mirror, camera mirror holder set and thermal camera. The research group of OMNIVIEWS project from Czech Technical University developed MATLAB software for designing omnidirectional mirror [9]. Utilizing the MATLAB software, omnidirectional hyperbolic mirror can be designed by inputting some parameters specify the mirror dimension. The details of the mirror design can be found in [13]. The coordinates generated using MATLAB software as well as mechanical drawing using Autocad was provided to precision engineering company to fabricate/custom made the hyperbolic mirror. The hyperbolic mirror is milling by using aluminum bar and chrome plating with chromium. Chromium is selected because of its lustrous (good in IR reflection), high corrosion resistance, high melting point and hardness. The camera mirror holder is designed and custom made with aluminum material as shown in Fig. 2.

The thermal camera used in the proposed omnidirectional thermal imaging system is an affordable and accurate temperature measurement model: ThermoVision A-20M manufactured by FLIR SYSTEM [10]. The thermal camera has a temperature sensitivity of 0.10 in a range from

-20°C to 900°C and it can capture thermal image with fine resolution up to 320 X 240 pixels offering more than 76,000 individual measurement points per image at a refresh rate of 50/60 Hz. For fast image and data transfer of real-time fully radiometric 16-bit images, an IEEE-1394 FireWire digital output can be selected. For network and/or multiple camera installations, Ethernet connectivity is also available. Each A-20M can be equipped with its own unique URL allowing it to be addressed independently via its Ethernet connection and it can be linked together with router to form a network. Therefore, it is best outfitted for machine condition monitoring system in a big factory site.

Log-polar mapper applying log polar mapping techniques [2]. It unwraps the captured omnidirectional thermal image into panoramic form, provided observer or image processing tools



FIGURE 2: Overall fabricated omnidirectional thermal imaging system model.

a wide angle of view. Another merit of log-polar image representation is that it has data compression manner. Log-polar mapping has been used in [13] for unwarping omnidirectional thermal image for machine condition monitoring. Detailed discussion on log-polar mapping will be provided in Section 3. Image partitioner is used for partitioned the input thermal image into S-partitioned sections, provided that the input thermal image consisted of S machines to be monitored. Each partitioned section consists of one machine to be monitored. An example of a panoramic thermal image with S=3 partition sections is shown in Fig. 3.

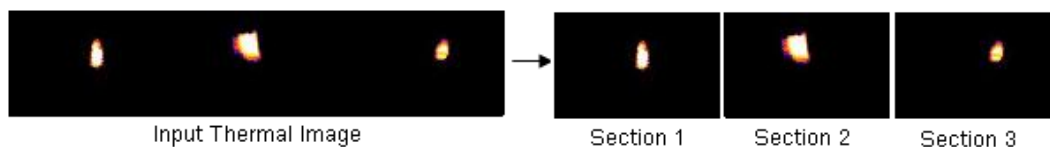


FIGURE 3: Example of a panoramic thermal image with S=3 partition sections.

Quaternion based thermal image correlator [5] is used to obtain correlation plane for each correlated input thermal image captured lively with reference images of all possible machines conditions stored in a database to calculate out some classification characteristics such as the real to complex ratio of the discrete quaternion correlation (DQCR) output, p -value and the peak-to-sidelobe ratio, PSR. These classification characteristics will later input to the max-product fuzzy neural network classifier to perform classification. Detailed discussion on quaternion based thermal image correlation will be performed in Section 4.

The max-product fuzzy neural network classifier [5] is first applied to train for an accurate classification with the weight (w) obtained from training reference images of all possible machine conditions stored in the database. During application, the PSR and ρ -value output from quaternion based thermal image correlator are first fuzzified into Gaussian membership function. Next, product value is calculated based on the multiplication of PSR in Gaussian membership value with ρ -value in Gaussian membership value. The product values are stored in an array and multiply with the weight (w). Max-composition is performing on the output based on two sets of fuzzy IF-THEN rules, and defuzzification is performed to classify each machine's condition under monitoring. Detailed discussion on max-product fuzzy neural network classifier will be given in Section 5.

3. LOG-POLAR MAPPING

Log-polar geometry or log-polar transform in short, is an example of foveated or space-variant image representation used in the active vision systems motivated by human visual system [11]. It is a spatially-variant image representation in which pixel separation increases linearly with distance from a central point [12]. It provides a way of concentrating computational resources on regions of interest, whilst retaining low-resolution information from a wider field of view. One merit of this kind of image mapping technique is data reduction. Foveal image representations like this are most useful in the context of active vision system where the densely sampled central region can be directed to pick up the most salient information. Mammals especially human eyes are very roughly organized in this way.

In the software conversion of log-polar images, practitioners in pattern recognition usually named it as log-polar mapping [13]. According to [13], it is found out that the use of log-polar mapping having numbers of merits such as help solving rotation and scaling problems in quaternion based thermal image correlation. However, the main merit of log-polar mapping to be applied in this paper is that it can unwarped/convert an omnidirectional image into panoramic image, hence providing the observer and image processing tools a complete wide angle of view for the surveillance area's surroundings and preserving fine output image quality in a higher data compression manner.

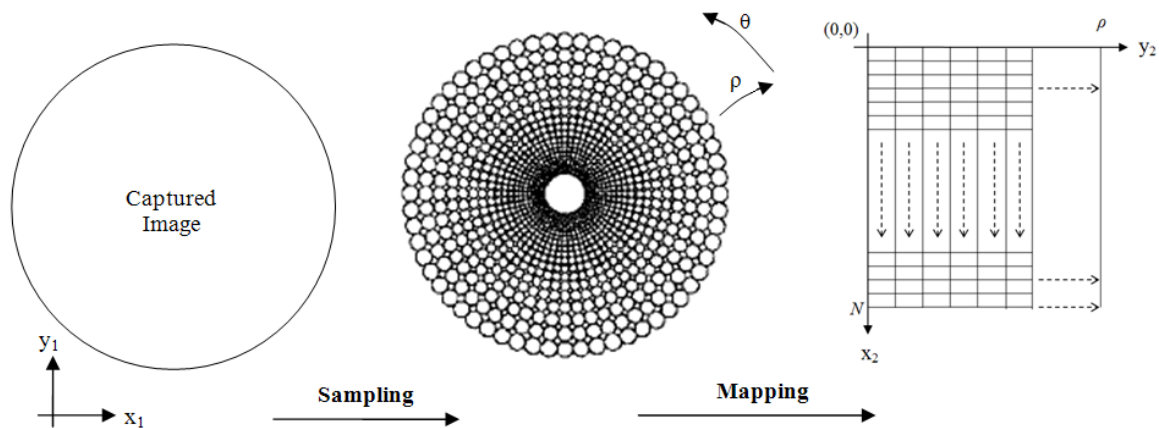


FIGURE 4 : A graphical view of log-polar mapping.

The spatially-variant grid that represents log-polar mapping is formed by i number of concentric circles with N samples over each concentric circle [11]. An example of a spatially-variant sampling grid is shown in Fig. 4.

The log-polar mapping use in this paper can be summarized as following [13]: Initially, omnidirectional thermal image is captured using the omnidirectional thermal imaging system as shown in Fig. 2. The geometry of the captured omnidirectional thermal image is in Cartesian form (x_1, y_1) . Next, the Cartesian omnidirectional thermal image is sampled by the spatially-variant grid into a log-polar form (ρ, θ) omnidirectional thermal image. After that, the log-polar omnidirectional

thermal image is unwarped into a panoramic thermal image (x_2, y_2) , another Cartesian form. Since the panoramic thermal image is in Cartesian form, subsequent image processing task will become much easier.

The centre of pixel for log-polar sampling is described by [2]:

$$\rho(x_1, y_1) = \log_{\lambda} \frac{R}{r_o} \quad (1)$$

$$\theta(x_1, y_1) = \frac{N_{\theta}}{2\pi} \arctan \frac{y_1}{x_1} \quad (2)$$

The centre of pixel for log-polar mapping is described by [2]:

$$x_2(\rho, \theta) = \lambda^{\rho} r_o \cos\left(\frac{2\pi\theta}{N_{\theta}}\right) \quad (3)$$

$$y_2(\rho, \theta) = \lambda^{\rho} r_o \sin\left(\frac{2\pi\theta}{N_{\theta}}\right) \quad (4)$$

where R is the distance between the given point and the center of mapping $= \sqrt{x_1^2 + y_1^2}$

r_o is a scaling factor which will define the size of the circle at $\rho(x_1, y_1) = 0$.

λ is the base of the algorithm,

$$\lambda = \frac{1 + \sin \frac{\pi}{N_{\theta}}}{1 - \sin \frac{\pi}{N_{\theta}}} \quad (5)$$

N_{θ} is the total number of pixels per ring in log-polar geometry. This value is assigned by user. For eg: if user assign $N_{\theta} = 5$, the ring is divided into 5 sectors or 5 pixels per ring. If $N_{\theta} = 100$, each ring is divided into 100 sectors or 100 pixels per ring. The higher the N_{θ} , the higher the resolution in θ -axis (angular). Refer [2] for detailed discussion on this issue.

The number of rings in the fovea region is given by [2]:

$$N_{fov} = \frac{\lambda}{\lambda - 1} \quad (6)$$

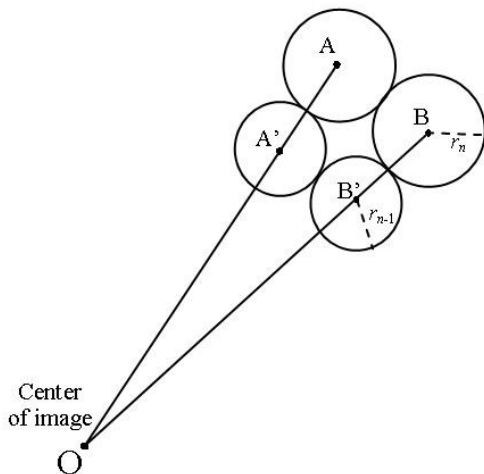


FIGURE 5 : Conventional circular sampling method for Log-Polar image.

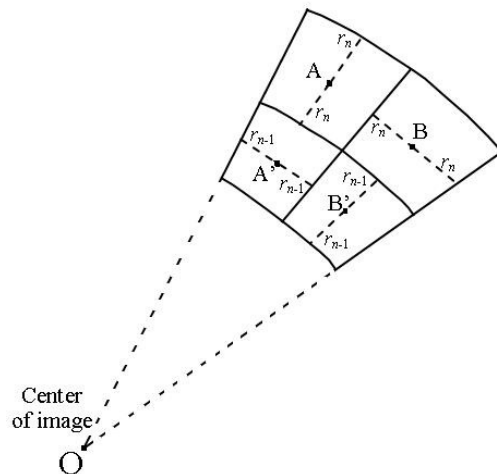


FIGURE 6 : Sector sampling method for log polar image.

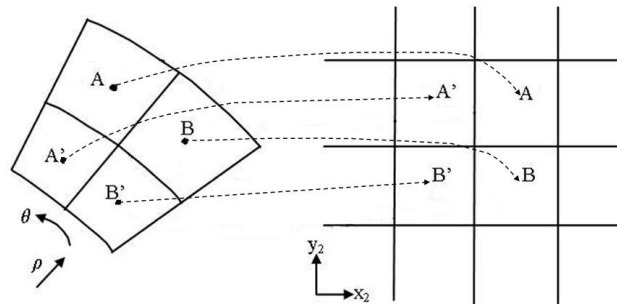


FIGURE 7 : Unwarping process

The number of rings must be an integer. Hence, the calculated N_{fov} is rounded to the closer integer value. To sample the Cartesian pixels (x_1, y_1) into log polar pixels (ρ, θ) , at each center point calculated using (1) and (2), the corresponding log-polar pixel (ρ_n, θ_n) is cover a region of Cartesian pixels with radius:

$$r_n = \lambda r_{n-1} \quad (7)$$

where $n=0, 1, \dots, N-1$. Fig. 5 shows the conventional circle sampling method of log-polar mapping [11, 14].

One of the demerits of using circle sampling is that certain region of Cartesian pixels outside sampling circle did not cover by any log-polar pixels. Therefore, some researchers [5, 15-17] had come out with sector sampling method as shown in Fig. 6, which could maximize the coverage of Cartesian pixels for each log polar pixel. The region of Cartesian pixels covers by an individual log-polar pixel will has the same color intensity follow the respective original Cartesian center sampling point.

During unwarping process, the (ρ, θ) pixels will map to each corresponding (x_2, y_2) pixels as shown in Fig. 7. The region of Cartesian pixels on the panoramic image (x_2, y_2) is covered by an individual log-polar pixel on the log-polar (ρ, θ) omnidirectional image. Therefore, the pixels in that specific region on the panoramic image (x_2, y_2) will have the same intensity with respect to the corresponding individual log-polar pixel.

4. QUATERNION BASED THERMAL IMAGE CORRELATOR

In this section, the algorithm of the quaternion based image correlator as proposed in [5] will be discussed.

4.1 Algorithm For Quaternion Based Thermal Image Correlator

The reference image after performing discrete quaternion Fourier transforms (DQFT) [4] is given by:

$$I(m, n) = I_R(m, n).i + I_G(m, n).j + I_B(m, n).k \quad (8)$$

where m, n are the pixel coordinates of the reference image. R, G, B parts of reference image are represented by $I_R(m, n)$, $I_G(m, n)$ and $I_B(m, n)$ respectively, and i, j, k are the imaginary parts of quaternion complex number [18], whereby the real part of it is set to zero. Similarly, $h_1(m, n)$ is used for representing input image. Then, output $b(m, n)$ can be produced to conclude whether the input image matches the reference image or not. If $h_1(m, n)$ is the space shift of the reference image:

$$h_1(m, n) = I(m - m_0, n - n_0) \quad (9)$$

then after some calculation,

$$\text{Max}(b_r(m, n)) = b_r(-m_0, n_0) \quad (10)$$

where $b_r(m, n)$ means the real part of $b(m, n)$ and

$$b_r(-m_0, n_0) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |I(m, n)|^2 \quad (11)$$

where M, N is the image x-axis, y-axis dimension. At the location $(-m_0, n_0)$, the multiplier of i, j, k -imaginary part of $b(-m_0, n_0)$ are equal to zero:

$$b_i(-m_0, n_0) = b_j(-m_0, n_0) = b_k(-m_0, n_0) = 0 \quad (12)$$

Hence, the process as below for thermal image correlation is followed [5]:

1.) Calculate energy of reference image $I(m, n)$:

$$E_I = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |I(m, n)|^2 \quad (13)$$

Then the reference image $I(m, n)$ and the input image $h_i(m, n)$ are normalized as:

$$I_a(m, n) = I(m, n) / \sqrt{E_I} \quad (14)$$

$$H_a(m, n) = h_i(m, n) / \sqrt{E_I} \quad (15)$$

2.) Calculate the output of discrete quaternion correlation (DQCR):

$$g_a(m, n) = \sum_{\tau=0}^{M-1} \sum_{\eta=0}^{N-1} I_a(\tau, \eta) \cdot \overline{H_a(\tau - m, \eta - n)} \quad (16)$$

where ' $\overline{\quad}$ ' means the quaternion conjugation operation and perform the space reverse operation:

$$g(m, n) = g_a(-m, -n) \quad (17)$$

3.) Perform inverse discrete quaternion Fourier Transform (IDQFT) on (17) to obtain the correlation plane $P(m, n)$.

4.) Search all the local peaks on the correlation plane and record the location of the local peaks as (m_s, n_s) .

5.) Then at all the location of local peaks (m_s, n_s) found in step 4, we calculate the real to complex value of the DQCR output is calculated:

$$p = \frac{|P_r(m_s, n_s)|}{|P_r(m_s, n_s)| + |P_i(m_s, n_s)| + |P_j(m_s, n_s)| + |P_k(m_s, n_s)|} \quad (18)$$

where $P_r(m_s, n_s)$ is the real part of $P(m_s, n_s)$. $P_i(m_s, n_s)$, $P_j(m_s, n_s)$ and $P_k(m_s, n_s)$ are the i, j, k - parts of $P(m_s, n_s)$ respectively. If $p \geq d_1$ and $c_1 < |P(m_s, n_s)| < c_2$, then it can be concluded that at location (m_s, n_s) , there is an object that has the same shape, size, color and brightness as the reference image. $d_1 < 1$, $c_1 < 1 < c_2$ and c_1, c_2 and d_1 are all with values near to 1. The

value of p decays faster with the color difference among the match image to the reference image.

Another classification characteristic proposed in quaternion based thermal image correlation by [5] is the peak-to-sidelobe ratio (PSR) which will be discussed in details as below. The quaternion based thermal image correlation involved 2 stages [5]: 1. Enrollment stage, and 2. Recognition stage. During the enrollment stage, one or multiple panoramic thermal images of each machine condition are acquired. These multiple reference images should have the variability in the color tones for different temperature conditions of the machines. The DQFT of the reference images are used to train fuzzy neural network and determine correlation filter for each possible machines' conditions. During recognition stage, omnidirectional thermal imaging system captures a live omnidirectional machines' thermal image, unwarps it into panoramic thermal image, and the DQFT of such image is correlated with the DQFT form of the reference images stored in the

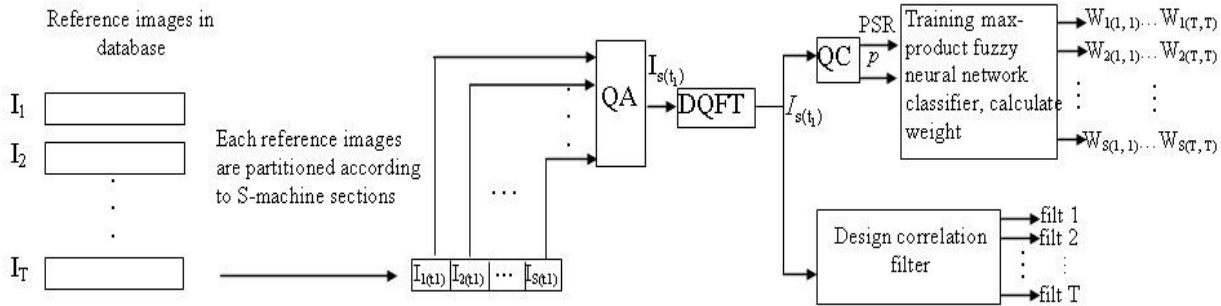


FIGURE 8 : Schematic of enrollment stage

database together with their corresponding filter coefficients, and the inverse DQFT of this product results in the correlation output for that filter.

A strong peak can be observed in the correlation output if the input image comes from imposter class. A method of measuring the peak sharpness is the peak-to-sidelobe ratio (PSR) which is defined as below [3, 5]:

$$PSR = \frac{peak - mean(sidelobe)}{\sigma(sidelobe)} \quad (19)$$

where $peak$ is the value of the peak on the correlation output plane. $sidelobe$ refers a fixed-sized surrounding area off the peak. $mean$ is the average value of the sidelobe region. σ is the standard deviation of the sidelobe region. Large PSR values indicate the better match of the input image and the corresponding reference image. Enrollment stage and recognition stage are discussed in details in next subsections.

4.2 Enrollment Stage

The schematic of enrollment stage for quaternion based thermal image correlator as proposed in [5] is shown in Fig. 8. To be applied in omnidirectional approach in this paper, during the enrollment stage, the reference panoramic thermal images for each possible machine's conditions in database are partitioned according to S machine sections. Each machine section consists of one single machine to be monitored. S is the total number of machines to be monitored. These partitioned reference images are then encoded into a two dimensional quaternion array (QA) as follows [5]:

$$I_{s(t_1)} = I_{sr(t_1)} + I_{sR(t_1)} \cdot i + I_{sG(t_1)} \cdot j + I_{sB(t_1)} \cdot k \quad (20)$$

where $t_1 = 1, 2, \dots, T$ represents the number of reference images, $I_{sr(t_1)}$ represents the real part of quaternion array of s -th machine section for reference image t_1 , $s = 1, 2, \dots, S$ represents the

number of partitioned machines' sections. $I_{sR(t_1)}$, $I_{sG(t_1)}$ and $I_{sB(t_1)}$ each represents the i , j , k -imaginary part of s -th machine section for reference image t_1 respectively.

The quaternion array in (20) is then performs discrete quaternion Fourier transform (DQFT) to transform the quaternion image to the quaternion frequency domain. A two-side form of DQFT has been proposed by Ell [19, 20] as follows:

$$I_{s(t_1)}(m, n) = \sum_{\tau=0}^{M-1} \sum_{\eta=0}^{N-1} e^{-\mu_1 2\pi(m\tau/M)} \cdot I_{s(t_1)}(\tau, \eta) \cdot e^{-\mu_2 2\pi(n\eta/N)} \quad (21)$$

where e is exponential term, μ_1 and μ_2 are two units pure quaternion (the quaternion unit with real part equal to zero) that are orthogonal to each other [21]:

$$\mu_1 = \mu_{1,i} \cdot i + \mu_{1,j} \cdot j + \mu_{1,k} \cdot k \quad (22)$$

$$\mu_2 = \mu_{2,i} \cdot i + \mu_{2,j} \cdot j + \mu_{2,k} \cdot k \quad (23)$$

$$\mu_{1,i}^2 + \mu_{1,j}^2 + \mu_{1,k}^2 = \mu_{2,i}^2 + \mu_{2,j}^2 + \mu_{2,k}^2 = 1 \quad (24)$$

(i.e. : $\mu_1^2 = \mu_2^2 = -1$)

$$\mu_{1,i} \cdot \mu_{2,i} + \mu_{1,j} \cdot \mu_{2,j} + \mu_{1,k} \cdot \mu_{2,k} = 0 \quad (25)$$

The output of DQFT, $I_{s(t_1)}$ is used to train the max-product fuzzy neural network classifier and design the correlation filter.

4.2.1 Quaternion Correlator (QC)

To train the max-product fuzzy neural network classifier, the output of the DQFT is first passed to a quaternion correlator (QC) as shown in Fig. 9 [5].

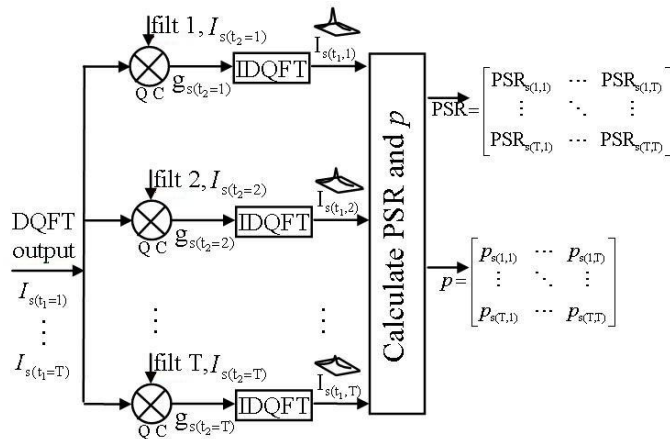


FIGURE 8 : Quaternion correlator (QC)

The function of the QC is summarized as below [5]: For DQFT output of s -th machine section, perform discrete quaternion correlation (DQCR) [22, 23] on reference image $I_{s(t_1)}$ with reference image $I_{s(t_2)}$ and multiply with corresponding filter coefficients (filt_(t2)):

$$g_{s(t_1, t_2)}(m, n) = \sum_{\tau=0}^{M-1} \sum_{\eta=0}^{N-1} I_{s(t_1)}(\tau, \eta) \cdot \overline{I_{s(t_2)}(\tau - m, \eta - n)} \cdot \text{filt}_{(t_2)} \quad (26)$$

where $t_1, t_2 = 1, 2, \dots, T$ are the number of reference image. After that, (26) is performing inverse DQFT to obtain the correlation plane function:

$$P_{s(t_1, t_2)}(m, n) = \frac{1}{4\pi^2} \sum_{\tau=0}^{M-1} \sum_{\eta=0}^{N-1} e^{-\mu_1 2\pi(m\tau/M)} \cdot g_{s(t_1, t_2)}(m, n) \cdot e^{-\mu_2 2\pi(m\eta/N)} \quad (27)$$

The correlation plane is a collection of correlation values, each one obtained by performing a pixel-by pixel comparison (inner product) of two images ($I_{s(t_1)}$ and $I_{s(t_2)}$). A sharp peak in the correlation plane indicate the similarity of $I_{s(t_1)}$ and $I_{s(t_2)}$, while the absence or lower of such peak indicate the dissimilarity of $I_{s(t_1)}$ and $I_{s(t_2)}$.

Calculate $p_{s(t_1, t_2)}$ and $PSR_{s(t_1, t_2)}$ from the correlation plane as in (27) using (18) and (19) respectively. $p_{s(t_1, t_2)}$ means p -values of reference image $I_{(t_1)}$ correlate on reference image $I_{(t_2)}$ in s -th machine section, while $PSR_{s(t_1, t_2)}$ means PSR values of reference image $I_{(t_1)}$ correlate on reference image $I_{(t_2)}$ in s -th machine section. These values are then feed into max-product fuzzy neural network classifier to perform training and calculate weight, which will be presented in section 5.

4.2.2 Correlation Filter Selection

According to research justification work done in [5], correlation filter outperforms conventional matched filters in filtering nonlinear image distortion (scale, rotation and pose invariant). Among the correlation filter, the minimum average correlation energy (MACE) filters [26] show good results in the field of automatic target recognition, face recognition and applications in biometric verification [6, 27]. MACE filters are using more than one reference image to synthesize a single filter template, therefore making its classification performance invariant to shift of the input image [25].

There are three types of MACE filters in general, namely: 1.) Conventional MACE filter [26], 2.) Unconstrained MACE (UMACE) filter [28] and 3.) Unconstrained optimal tradeoff synthetic discriminant filter (UOTSDF), all with the goal to produce sharp peaks that resemble two dimensional delta-type correlation outputs when the input image belongs to the authentic class and low peaks in imposter class. Conventional MACE filter [26] minimizes the average correlation energy of the reference images while constraining the correlation output at the origin to a specific value (usually 1), for each of the reference images. Lagrange multiplier is used for optimization, yielding:

$$\text{filt}_{\text{MACE}} = D^{-1} X(X'D^{-1}X)^{-1} c \quad (28)$$

This equation is the closed form solution to be the linear constrained quadratic minimization. D is diagonal matrix with the average power spectrum of the reference images placed as elements along diagonal of the matrix. X contains Fourier transform of the reference images lexicographically re-ordered and placed along each column. As an example, if there are T thermal reference images of size $282 \times 60 (=16920)$, then X will be a $16920 \times T$ matrix. X' is the matrix transpose of X . c is a column vector of length T with all entries equal to 1.

The second type of MACE filter is the unconstrained MACE (UMACE) filter [28]. Just like conventional MACE filter, UMACE filter also minimizes the average correlation energy of the reference images and maximizes the correlation output at the origin. The different between conventional MACE filter and UMACE filter is the optimization scheme. Conventional MACE filter is using Lagrange multiplier but as for UMACE filter, it is using Raleigh quotient which lead to the following equation:

$$\text{filt}_{\text{UMACE}} = D^{-1} m \quad (29)$$

where D is the diagonal matrix same as that in conventional MACE filter. m is a column vector containing the mean values of the Fourier transform of the reference images.

The third type of MACE filter is the unconstrained optimal tradeoff synthetic discriminant filter (UOTSDF) shown by Refreiger [29] and Kumar et al [30] has yielding good verification performance. The UOTSDF is by:

$$fil_{UOTSDF} = (\alpha D + \sqrt{1 - \alpha^2} C)^{-1} m \quad (30)$$

where D is a diagonal matrix with average power spectrum of the training image placed along the diagonal elements. m is a column vector containing the mean values of the Fourier transform of the reference images. C is the power spectral density of the noise. White noise spectrum is the dominant source in predicting the performance of a thermal imaging system [31]. It is cause by the fluctuation in the detector output. Other noise sources (total up as background noise) are not that significant and normally limited /filter out by internal filter of some advanced thermal imaging system. For most of the applications, a white noise power spectral density is for assumption, therefore C reduces to the identity matrix. According to the derivation work done in [30], to determine the OTSDF, the authors minimize the energy function which obtains: $\alpha^2 + \beta^2 + \gamma^2 + \delta^2 = 1$. In UOTSDF, the constant $\beta, \gamma, \delta \approx 0$. α term is typically set to be close to 1 to achieve good performance even in the presence of noise, however it also helps improve generalization to distortions outside the reference images.

The comparison work of the three correlation filters listed above is done in [5]. As a summary, UOTSDF is plan to extend into quaternion based thermal image correlator for the classification of

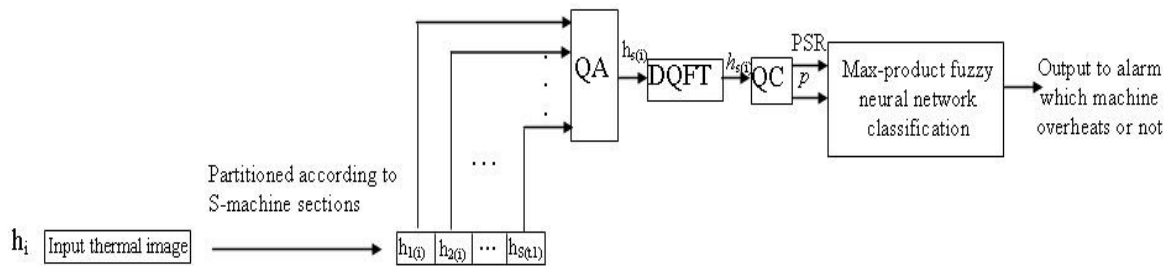


FIGURE 10 : Schematic of recognition stage

machine condition since it is less complicated in computational viewpoint than conventional MACE filter and achieve good performance.

4.3 Recognition Stage

The schematic of recognition stage for classification of machine condition by quaternion based thermal image correlator as proposed in [5] is shown in Fig. 10. To be applied in omnidirectional approach in this paper, during the recognition stage, live omnidirectional machines' image captured by omnidirectional thermal imaging system, unwarped into panoramic thermal image is first partitioned according to S machine sections. The partitioned image is then encoded into two dimensional quaternion array (QA) as follows [5]:

$$h_{s(i)} = h_{sr(i)} + h_{sR(i)} \cdot i + h_{sG(i)} \cdot j + h_{sB(i)} \cdot k \quad (31)$$

where i represents the input image, $h_{sr(i)}$ represents the real part of quaternion array of s-th machine section for input image i, $s = 1, 2, \dots, S$ represents the number of partitioned machines' sections. $h_{sR(i)}, h_{sG(i)}$ and $h_{sB(i)}$ each represents the i, j, k -imaginary part of s-th machine section for input image i respectively.

The quaternion array in (31) is then performing DQFT to transforms the quaternion image to the quaternion frequency domain. A two-side form of DQFT is used [5]:

$$h_{s(i)}(m, n) = \sum_{\tau=0}^{M-1} \sum_{\eta=0}^{N-1} e^{-\mu_1 2\pi(m\tau/M)} \cdot h_{s(i)}(\tau, \eta) \cdot e^{-\mu_2 2\pi(n\eta/N)} \quad (32)$$

where e is exponential term, μ_1 and μ_2 are two units pure quaternion as shown in (22) and (23) respectively. The output of the DQFT, $h_{s(i)}$ is cross correlated with every quaternion correlation filter in the database using the quaternion correlator (QC) just as the one shown in Fig. 8, but the DQFT output is now $h_{s(i)}$. In QC, performs quaternion correlation is performs on $h_{s(i)}$ with reference images $I_{s(t_2)}$ from database, and multiply with corresponding filter coefficients ($filt_{(t_2)}$) [5]:

$$g_{s(i, t_2)}(m, n) = \sum_{\tau=0}^{M-1} \sum_{\eta=0}^{N-1} h_{s(i)} \cdot \overline{I_{s(t_2)}(\tau - m, \eta - n)} \cdot filt_{(t_2)} \quad (33)$$

After that, (33) is performing inverse DQFT to obtain the correlation plane function [5]:

$$P_{s(i, t_2)}(m, n) = \frac{1}{4\pi^2} \sum_{\tau=0}^{M-1} \sum_{\eta=0}^{N-1} e^{-\mu_1 2\pi(m\tau/M)} \cdot g_{s(i, t_2)}(m, n) \cdot e^{-\mu_2 2\pi(n\eta/N)} \quad (34)$$

Calculate $p_{s(i, t_2)}$ and $PSR_{s(i, t_2)}$ from the correlation plane as in (34) using (18) and (19) respectively. $p_{s(i, t_2)}$ means p -values of input image $h_{(i)}$ correlate on reference image $I_{(t_2)}$ in s -th machine section, while $PSR_{s(i, t_2)}$ means PSR values of input image $h_{(i)}$ correlate on reference image $I_{(t_2)}$ in s -th machine section. These values are then feed into max-product fuzzy neural network classifier to perform classification for machines' conditions, which will be presented in section 5.

5. MAX-PRODUCT FUZZY NEURAL NETWORK CLASSIFIER

Fuzzy logic is a problem solving system which is capable of dealing with approximate reasoning. Fuzzy logic provides high level of abstract through process which can appropriately handle the uncertainty in linguistic semantics, model expert heuristics and provide requisite high level organizing principles [32]. On the other hand, neural network is a computational biological network that can provides self organizing substrates for low level representation of information with adaptation capabilities. Both fuzzy logic and neural network are complimentary technologies and these two approaches is plausible to combine in the design of classification systems. Such integrated system is terms as fuzzy neural network classifier [32].

Many types of fuzzy neural network classifiers available in literature [33-36], and many fuzzy neural networks had been shown interest in applying max-min composition as functional basis [37-39]. However, Leotamonphong and Fang in their research work [40] mentioned that the max-min composition is "suitable only when a system allows no compensability among the elements of a solution vector". They proposed to use max-product composition in fuzzy neural network rather than max-min composition. Research work done by Bourke and fisher in [41] also commented that the max-product composition gives better results than the traditional max-min operator. Subsequently, many efficient learning algorithms have been studied by others [42, 43] using the max-product composition.

Fuzzy neural network classifier using max-product composition for thermal image classification has been proposed in [5] for fixed angle machine condition monitoring system. In this paper, it will be once again repurposed for omnidirectional approached machine condition monitoring system.

5.1 Define 2 Classes, Namely: Overheat Class and Non-overheat Class

The reference images for all possible machines' conditions are captured, unwarped into panoramic form and stored in a database. Each of these reference images will be assigned with a unique number start from 1 till T, where T is the total number of reference images. These reference images are interpreted by an operator (human observer), the overall description of

which could be called the ‘Operator perceived activity’ (OPA)[44]. The operator will comments on each of the reference images and classified it into either overheat class or non-overheat class by storing the unique number of the reference images according to the classes respectively.

5.2 Training Max-product Fuzzy Neural Network Classifier

The max-product fuzzy neural network classifier is training with 4 steps [5]:

- 1.) $PSR_{s(t_1,t_2)}$ and $p_{s(t_1,t_2)}$ output from the quaternion correlator of the enrollment stage are fuzzified through the activation functions (Gaussian membership function):

$$G_{PSR_{s(t_1,t_2)}} = \exp\left[\frac{-(PSR_{s(t_1,t_2)} - 1)^2}{\sigma^2}\right] \quad (35)$$

$$G_{p_{s(t_1,t_2)}} = \exp\left[\frac{-(p_{s(t_1,t_2)} - 1)^2}{\sigma^2}\right] \quad (36)$$

where σ is the smoothing factor, that is the deviation of the Gaussian functions.

- 2.) Calculate the product value for s-th machine section of the fuzzy neural network classifier at each correlated images:

$$G_{s(t_1,t_2)} = G_{PSR_{s(t_1,t_2)}} \times G_{p_{s(t_1,t_2)}} \quad (37)$$

- 3.) Gather and store the product values in an array:

$$X_{s\text{ training}} = \begin{bmatrix} G_{s(1,1)} & G_{s(1,2)} & \cdots & G_{s(1,T)} \\ G_{s(2,1)} & G_{s(2,2)} & \cdots & G_{s(2,T)} \\ \vdots & \vdots & \ddots & \vdots \\ G_{s(T,1)} & G_{s(T,2)} & \cdots & G_{s(T,T)} \end{bmatrix} \quad (38)$$

- 4.) The output will set so that it will output 1 if it is authentic class and 0 if it is imposter class, and it is in an array Y_{identity} , whereby it is an identity matrix of dimension $T \times T$. To calculate the weight w for s-th machine section, the equation is:

$$w_s = X_{\text{straining}}^{-1} Y_{\text{identity}} \quad (39)$$

5.3 Max-product Fuzzy Neural Network Classification

The max-product fuzzy neural network classification is with 7 steps [5]:

- 1.) $PSR_{s(i,t_2)}$ and $p_{s(i,t_2)}$ output from the quaternion correlator of the recognition stage are fuzzified through the activation functions (Gaussian membership function):

$$G_{PSR_{s(i,t_2)}} = \exp\left[\frac{-(PSR_{s(i,t_2)} - 1)^2}{\sigma^2}\right] \quad (40)$$

$$G_{p_{s(i,t_2)}} = \exp\left[\frac{-(p_{s(i,t_2)} - 1)^2}{\sigma^2}\right] \quad (41)$$

- 2.) Calculate the product value for s-th machine section of the fuzzy neural network classifier at input image on the training images in database:

$$G_{s(i,t_2)} = G_{PSR_{s(i,t_2)}} \times G_{p_{s(i,t_2)}} \quad (42)$$

3.) Gather and store the product values in an array:

$$X_{s \text{ classification}} = [G_{s(i,1)} \quad G_{s(i,2)} \quad \cdots \quad G_{s(i,T)}] \quad (43)$$

4.) Obtain the classification outcomes for each machine condition in the s-th section by multiply (43) with the weight trained at (39):

$$Y_{s \text{ classification}} = X_{s \text{ classification}} \times w_s \quad (44)$$

5.) Classify the input machine condition with the class of machine condition it belongs to by using max composition:

$$\text{Class}_s = \max\{Y_{s \text{ classification}}\} \quad (45)$$

6.) Determine which element in $Y_{s \text{ classification}}$ matrix match with Class_s :

$$\psi = \text{the position number of element in } Y_{s \text{ classification}} \text{ matrix which has the equal value with } \text{Class}_s. \quad (46)$$

ψ is corresponds to the assigned number of reference image in database.

7.) Based on two sets of fuzzy IF-THEN rules, perform defuzzification:

$$R_s^1: \text{ IF } \psi \text{ is match with the number stored in overheat class of s-th machine, THEN alarm: 'machine s overheat'.} \quad (47)$$

$$R_s^2: \text{ IF } \psi \text{ is match with the number stored in non-overheat class of s-th machine, THEN alarm: 'machine s function properly'.} \quad (48)$$

6. EXPERIMENTAL RESULTS

In this section, the application of log-polar mapper, quaternion based thermal image correlator together with max-product fuzzy neural network classifier is briefly illustrate for omnidirectional machine condition monitoring system. Here, some experiments are use to prove the algorithms introduced in section 3, 4 and 5.

6.1 Database of Reference Thermal Images for All Possible Machines' Condition.

A database with thermal images collected at the Applied Mecahnics Lab in Faculty of Engineering and Technology, Multimedia University is use to test the proposed quaternion based omnidirectional machine condition monitoring system. The database consists of panoramic thermal images unwrap from omnidirectional thermal images captured from three functioning machines by the use of omnidirectional thermal imaging system as shown in Fig. 2. A digital image captured with digital camera on the site is shown in Fig. 11, and a thermal image is also captured by the omnidirectional thermal camera set as shown in Fig. 12 at the same position based on all the functioning machines are in overheat condition. Comparing Fig. 11 and Fig. 12, observed that the material (mirror) good in IR reflection not necessary good in human visual range reflection. The corresponding panoramic form unwrapped using log-polar mapping of Fig. 11 and Fig. 12 are shown in Fig. 13 and Fig. 14 respectively. In Fig. 13, machine A (leftmost one) and machine C (rightmost one) are vibro test machines with same model and same specifications, while machine B (center one) is a fatigue test machine. Three machines are considered to be overheating when their motors' temperature achieves 90°C.

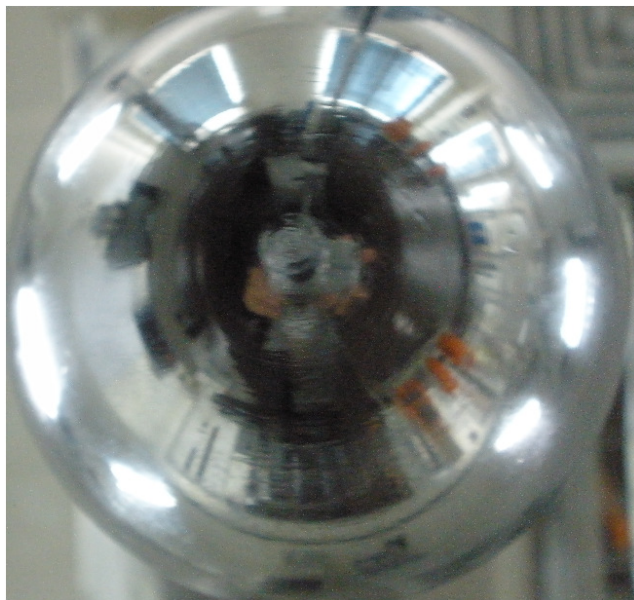


FIGURE 11 : Digital color form on site image site image



FIGURE 12 : Thermal on site image



FIGURE 13 : Panoramic form of Fig. 11.



FIGURE 14 : Panoramic form of Fig. 12.

The thermal images captured using FLIR ThermoVision A20M is with 320x240 display resolution pixels. Log-polar mapper will unwarped it into panoramic images with 282x60 display resolution pixels. The log-polar mapping process is by 4.54:1 reduction mapping scale, with almost four and a half fold data compression compare to original omnidirectional thermal image. Each machine section later partition into 94x60 pixels, as shown in Fig. 3. The database has $T = 30$ reference images, each with dimension 282 horizontal pixels \times 60 vertical pixels of varying possible machines' conditions (temperature level with different color tones ranging from black, blue, purple, red, orange, yellow, light yellow to white) and can be divided into 8 major outcomes, namely: 1.) All machines function properly (none of the machines overheat), 2.) machine A overheat, 3.) machine B overheat, 4.) machine C overheat, 5.) machine A and machine B overheat, 6.) machine A and machine C overheat, 7.) machine B and machine C overheat, 8.) machine A, B, and C overheat.

6.2 Unconstrained Optimal Tradeoff Synthetic Discriminant Filter (UOTSDF) Used in Quaternion Based Thermal Image Correlator

The 30 reference images stored in database is use to synthesize a single UOTSDF using (30). D , m are calculated from the reference images and C is set as an identity matrix of dimension 30×30 and α set to 1. These values are substituted into (30) to calculate out the filter's coefficients. In enrollment stage, for each filter line as in Fig. 8, cross correlations are performed on all the DQFT form of reference images in database ($I_{s(t_1)}$) with the DQFT form of reference images in database as well ($I_{s(t_2)}$), and multiply the output value with corresponding filter coefficients respectively, where $t_1, t_2 = 1, 2, \dots, 30; s = 1, 2, 3$. In recognition stage, for each filter line, cross correlation are also performed on the DQFT form of input image ($h_{s(i)}$) with the DQFT form of reference images in database ($I_{s(t_2)}$) and multiply the output value with corresponding filter coefficients respectively. For authentic case (good match in between two images), the correlation plane should have sharp peaks and it should not exhibit such strong peaks for

imposter case (bad or no match in between two images). These two cases will be investigated below:

Authentic case: Fig. 16a-16c show the samples correlation plane for input thermal image (for every machine is overheat, as in Fig. 15) matching with one of the reference image of overheat class in the database, for section machine A, section machine B and section machine C respectively. Since the three pairs of images are in good match, their correlation planes are having smooth and sharp peaks.

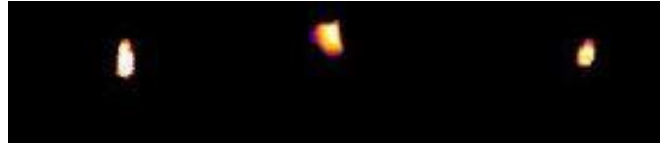


FIGURE 15 : Sample input thermal image (all machines overheat)

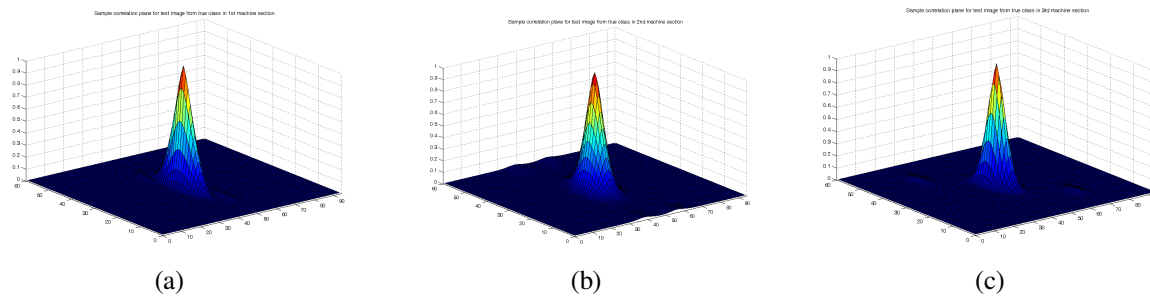


FIGURE 16 : Samples correlation plane for input thermal image (for every machine is overheat) matching with one of the reference image of overheat class for both all the machines in the database (authentic case) a.) section machine A, b.) section machine B, c.) section machine C.

Imposter case: Fig. 18a-18c show the samples correlation plane for input thermal image (for every machine is not –overheat as in Fig. 17) matching with one of the reference image of overheat class in the database (as in Fig. 12), for section machine A, section machine B and section machine C respectively. Since the three pairs of images are not in good match, their correlation planes are having no sharp peak at all.

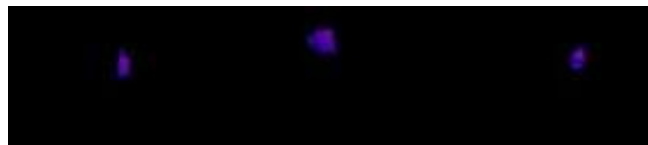


FIGURE 17 : Sample input thermal image (all machines not overheat)

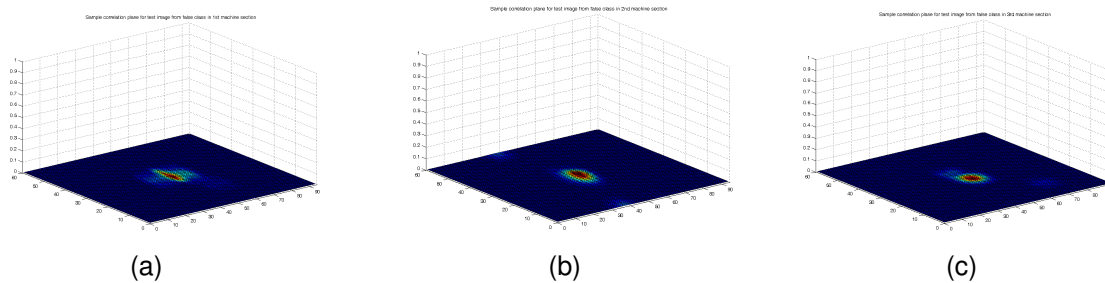


FIGURE 18 : Samples correlation plane for input thermal image (for every machine is not-overheat) matching with one of the reference image of overheat class for both all the machines in the database (imposter case) a.) section machine A, b.) section machine B, c.) section machine C.

Table 1 shows the PSR and p -value for both authentic and imposter case as in Fig. 16 and Fig 18 for section machine A, B and C. Note that the sharp correlation peak resulting in large normalized PSR and p -value in authentic case of section machine A, B and C, whereas small PSR and p -value exhibiting in the imposter case of section machine A, B and C.

TABLE 1: Normalized PSR and p -values for both authentic and imposter case.

Authentic case	Normalized PSR	Normalized p -value	Imposter case	Normalized PSR	Normalized p -value
Section mac. A	0.9820	0.9886	Section mac. A	0.0453	0.0363
Section mac. B	0.9437	0.9836	Section mac. B	0.0365	0.0315
Section mac. C	1.0000	0.9947	Section mac. C	0.0489	0.0375

6.3 Efficiency of the Quaternion Based Machine Condition Monitoring System

The quaternion based omnidirectional machine monitoring system was evaluated with respect to the thermal images captured live, unwarped into panoramic form and displayed on monitor screen as interpreted by a operator (human observer) the overall description of which could be called the 'operator perceived activity' (OPA) [45]. The operator will comments on the unwarped panoramic images captured by the omnidirectional thermal imaging system, whether any of the machines are overheat or not and compare with that classified by the proposed machine condition monitoring system. The system was evaluated for 10,000 samples images captured by the omnidirectional thermal imaging system for monitoring the functioning machines as in Fig. 11. Among the total 10,000 samples images, 9,702 were tracked perfectly (output machines' conditions agreed by both observer and the machine condition monitoring system), i.e. an overall accuracy of 97.02%.

7. CONCLUSIONS

This paper presented an omnidirectional machine condition monitoring system capable of monitoring machine condition in a wide area coverage using minimum hardware manner, whereby the machines surrounded in an omnidirectional (360°) view can be monitored by using a single thermal camera and a custom made hyperbolic IR mirror. The proposed machine condition monitoring system also using log-polar mapper unwarping 320x240 omnidirectional thermal images into 282x60 panoramic image, providing observer or image processing tools a wide angle of view and with data compression upto 4.5 folds. Therefore, log-polar mapping helps reduces the computation time in image processing and memory storage needs. Quaternion thermal image correlation method deals with color thermal images without converting them into gray-scale images. This can better preserved important color information. Max-product fuzzy neural network is a high level framework for approximate reasoning, best suit to be used in classification of machine conditioning. Experimental results show that the proposed machine condition monitoring system is with accuracy as high as 97 %. The experimental results show that the apply of PSR and p -value give higher accuracy in tracking thermal condition for the proposed quaternion based omnidirectional machine condition monitoring system. In future, collaboration is plan seek for implementing the proposed quaternion based omnidirectional machine condition monitoring system in fossil power plant monitoring. The omnidirectional scenes in a site of turbine room, steam pipework system, boilers, high pressure by-pass drain valve etc within a fossil power plant can be monitored effectively with such system. A trespasser detection algorithm is also plan to develop for trespasser detection purpose. All these topics will be addressed in future work.

REFERENCES

- [1] W. K. Wong, P. N. Tan, C. K. Loo and W.S. Lim, "An Effective Surveillance System Using Thermal Camera", 2009 International Conference on Signal Acquisition and Processing (ICSAP 2009), 3-5, Apr 2009, Kuala Lumpur, Malaysia: 13-17.
- [2] F. Berton, A brief introduction to log-polar mapping, Technical report, LIRA-Lab, University of Genova, (Feb 2006).

- [3] C. Xie, M. Savvides and B.V.K. Vijaya Kumar, "Quaternion correlation filters for face recognition in wavelet domain", Int. Conf. on Accoustic, Speech and Signal Processing (ICASSP 2005):1185- 1188.
- [4] S. C. Pei, J. J. Ding and J. Chang, "Color pattern recognition by quaternion correlation", Proc. of Int. Conf. on Image Processing, Vol.1, (2001) : 894-897.
- [5] W. K. Wong, C. K. Loo, W. S. Lim, P. N. Tan, "Thermal Condition Monitoring System Using Log-Polar Mapping, Quaternion Correlation and Max-Product Fuzzy Neural Network Classification", Elsevier, Neurocomputing (in press).
- [6] B.V.K. Vijaya Kumar, M. Savvides, K. Venkataramani and C. Xie, "Spatial frequency domain image processing for biometric recognition", Proc. Of Int. Conf. on Image Processing, Vol.1, (2002) p.p. 153-156.
- [7] R.O. Duda, P.E. Hart and D.G. Stork, Pattern Classification, 2nd Ed. Wiley, N.Y, 2001.
- [8] R.K. Brouwer, "A fuzzy threshold max-product unit, with learning algorithm, for classification of pattern vectors", Proc. of the VI Brazillian Symp. On Neural Networks, (Jan 2000) p.p. 208-212.
- [9] S. Gatcher, "Mirror Design for an Omnidirectional Camera with a Uniform Cylindrical Projection When Using the SVAVISCAS Sensor", Research Reports of CMP, OMNIVIEWS Project, Czech technical University in Prague, No. 3, 2001. Redirected from: <http://cmp.felk.cvut.cz/projects/omniviews/>
- [10] <http://www.flirthermography.com>
- [11] H. Araujo, J. M. Dias, "An Introduction To The Log-polar Mapping", Proceedings of 2nd Workshop on Cybernetic Vision, 1996, p.p. 139-144.
- [12] C. F. R. Weiman and G. Chaikin, "Logarithmic Spiral Grids For Image Processing And Display", Computer Graphics and Image Processing, Vol. 11, 1979, p.p. 197-226.
- [13] W. K. Wong, P. N. Tan, C. K. Loo and W. S. Lim , "Machine Condition Monitoring Using Omnidirectional Thermal Imaging System", IEEE International Conference on Signal & Image Processing Applications 18-19, November 2009, Kuala Lumpur, Malaysia, Paper No. 151, p.p. 1-6.
- [14] R. Wodnicki, G.W. Roberts, and M.D. Levine, "A foveated image sensor in standard CMOS technology", Custom Integrated Circuits Conf. Santa Clara, May 1995, p.p. 357-360.
- [15] F. Jurie, "A new log-polar mapping for space variant imaging: Application to face detection and tracking", Pattern Recognition, Elsevier Science, 32:55, 1999, p.p. 865-875.
- [16] M. Bolduc and M.D. Levine, "A review of biologically-motivated space variant data reduction models for robotic vision", Computer Vision and Image Understanding, Vol. 69, No. 2, (February 1998) p.p. 170-184.
- [17] C.G. Ho, R.C.D. Young and C.R. Chatwin, "Sensor geometry and sampling methods for space variant image processing", Pattern Analysis and Application, Springer Verlag, (2002) p.p. 369-384.
- [18] W. R. Hamilton, Elements of Quaternions, London, U.K.: Longmans, Green (1866).

- [19] T.A. Ell, "Quaternion-Fourier transforms for analysis of two-dimensional linear time-invariant partial differential systems", Proc. of 32nd Conf. Decision Contr., (Dec 1993) p.p. 1830-1841.
- [20] T.A. Ell, "Hypercomplex spectral transforms", PhD dissertation, Univ. Minnesota, Minneapolis, 1992.
- [21] S.C. Pei, J.J. Ding and J.H. Chang, "Efficient implementation of quaternion Fourier transform convolution and correlation by 2-D Complex FFT", IEEE Trans. on Signal Processing, Vol. 49, No. 11, (Nov 2001) p.p. 2783-2797.
- [22] S.J. Sangwine and T.A. Ell, "Hypercomplex auto- and cross-correlation of colour images", Proc. of Int. Conf. on Image Processing, (ICIP 1999) p.p. 319-323.
- [23] T.A. Ell and S.J. Sangwine, "Colour –sensitive edge detection using hypercomplex filters", (EUSIPCO 2000) p.p. 151-154.
- [24] A Vanderlugt, "Signal detection by complex spatial filtering", IEEE Trans. Inf. Theory, Vol. 10, (1964) p.p.139-145.
- [25] M. Saviddes, K. Venkataramani and B.V.K. Vijaya Kumar, "Incremental updating of advanced correlation filters for biometric authentication systems", Proc. of Int. Conf. on Multimedia and Expo, Vol. 3 (ICME 2003) p.p. 229-232.
- [26] A. Mahalanobis, B.V.K. Vijaya Kumar and D. Casasent, "Minimum average correlation energy filters", Applied Optics, Vol. 26, (1987) p.p. 3633-3640.
- [27] M. Savvides, B.V.K. Vijaya Kumar and P. Khosla, "Face verification using correlations filters", Procs of 3rd IEEE Automatic Identification Advanced Technologies, Tarrytown, N.Y., (2002) p.p. 56-61.
- [28] A. Mahalanobis, B.V.K. Vijaya Kumar, S.R.F. Sims and J.F. Epperson, "Unconstrained correlation filters", Applied Optics, Vol. 33, (1994) p.p. 3751-3759.
- [29] P. Refreiger, "Filter design for optical pattern recognition: multi-criteria optimization approach", Optics Letters, Vol. 15, (1990) p.p. 854-856.
- [30] B.V.K. Kumar, D.W. Carlson, and A. Mahalanobis, "Optimal trade-off synthetic discriminant function filters for arbitrary devices", Optics Letters, Vol. 19, No. 19, (1994) p.p. 1556-1558.
- [31] H. V. Kennedy, "Modeling noise in thermal imaging systems", Proc. of SPIE, Vol. 1969, (1993), p.p. 66-70.
- [32] S. Kumar, Neural Networks: A Classroom Approach, McGraw Hill, Int. Ed., 2004.
- [33] J.J. Buckley and Y. Hayashi, "Fuzzy neural networks: A survey", Fuzzy Sets and Systems, 66, (1994) p.p. 1-13.
- [34] C.T. Lin and C.S.G. Lee, Neural Fuzzy Systems: A Neuro-Fuzzy Synergism to Intelligent Systems, Prentice Hall, Upper Saddle River, N.J., 1996.
- [35] D. Nauck, F. Klawonn and R. Kurse, Foundations of Neuro-Fuzzy Systems, Wiley, Chichester, U.K., 1997.
- [36] S.K. Pal and S. Mitra, Neuro-Fuzzy Pattern Recognition: Methods in Soft Computing, Wiley, Chichester, U.K., 1999.
- [37] R. Ostermark, "A Fuzzy Neural Network Algorithm for Multigroup Classification", Elsevier Science, Fuzzy Sets and Systems, 105, (1999) p.p. 113-122.

- [38] H.K. Kwan and Y. Cai, "A Fuzzy Neural Network and its Application to Pattern Recognition", IEEE Trans. on Fuzzy Systems, 2(3), (1997) p.p. 185-193.
- [39] G.Z. Li and S.C. Fang, "Solving interval-valued fuzzy relation equations", IEEE Trans. on Fuzzy Systems, Vol. 6, No. 2, (May 1998) p.p. 321-324.
- [40] J. Leotamonphong and S. Fang, "An efficient solution procedure for fuzzy relation equations with max product composition", IEEE Trans. on Fuzzy Systems, Vol. 7, No. 4, (Aug 1999) p.p. 441-445.
- [41] M.M. Bourke and D.G. Fisher, "A predictive fuzzy relational controller", Proc. of the Fifth Int. Conf. on Fuzzy Systems, (1996) p.p. 1464-1470.
- [42] M.M. Bourke and D.G. Fisher, "Solution algorithms for fuzzy relational equations with max-product composition", Fuzzy Sets Systems, Vol. 94, (1998) p.p. 61-69.
- [43] P. Xiao and Y. Yu, "Efficient learning algorithm for fuzzy max-product associative memory networks", SPIE, Vol. 3077, (1997), p.p. 388-395.
- [44] J. Owens, A. Hunter and E. Fletcher, "A Fast Model-Free Morphology-Based Object Tracking Algorithm", British Machine Vision Conference, p.p. 767-776, 2002.

Automatic Threshold based Liver Lesion Segmentation in Abdominal 2D-CT Images

Asmita A. Moghe

*Department of IT University Institute of Technology
(Rajiv Gandhi Proudyogiki Vishwavidyalaya)
Bhopal, 462036, India*

aamoghe@gmail.com

Dr. Jyoti Singhai

*Department of Electronics & Communication
Maulana Azad National Institute of Technology
Bhopal, 462051, India*

j.singhai@gmail.com

Dr. S.C Shrivastava

*Department of Electronics & Communication
Maulana Azad National Institute of Technology
Bhopal, 462051, India*

scs_manit@yahoo.com

Abstract

Liver lesion segmentation using single threshold in 2D abdominal CT images proves insufficient. The variations in gray level between liver and liver lesion, presence of similar gray levels in adjoining liver regions and type of lesion may vary from person to person. Thus, with threshold based segmentation, choice of appropriate thresholds for each case becomes a crucial task. An automatic threshold based liver lesion segmentation method for 2D abdominal CT pre contrast and post contrast image is proposed in this paper. The two thresholds, Lower Threshold and Higher Threshold are determined from statistical moments and texture measures. In pre contrast images, gray level difference in liver and liver lesion is very feeble as compared to post contrast images, which makes segmentation of lesion difficult. Proposed method is able to determine the accurate lesion boundaries in pre-contrast images also. It is able to segment lesions of various types and sizes in both pre contrast and post contrast images and also improves radiological analysis and diagnosis. Algorithm is tested on various cases and four peculiar cases are discussed in detail to evaluate the performance of algorithm.

Keywords: Segmentation, lesion, Thresholding.

1. INTRODUCTION

Imaging modalities like Ultrasound, MRI (Magnetic Resonance Imaging), CT (Computed Tomography) and PET (Positron Emission Tomography) are widely used techniques for liver lesion diagnosis. However, CT is ubiquitously available and preferred imaging modality among clinical practitioners. Abdominal CT images are segmented to determine different organs, the lesions in the abdominal organs and to provide 3D rendering of these organs.

The liver is the largest organ in the body. The contour and shape of the liver vary according to the patient habitus and lie [1]. Its shape is also influenced by the lateral segment of the left lobe and the length of the right lobe. The ribs cover the greater part of the right lobe and usually a small part of the right lobe is in contact with the abdominal wall. The right lobe is the largest of the four lobes of the liver and exceeds the left lobe by a ratio of 6:1. In an abdominal CT image of a healthy subject, the liver has uniform gray level. In the subjects with liver lesions, the gray level of lesion is different from the gray level of liver. This difference is better enhanced in corresponding contrast enhanced images. In clinical practice radiologists assess the lesion based on pre

contrast images and the post contrast (contrast enhanced) images obtained for the same patient after injection of contrast either for diagnosis or for use during pre operative and/or intra operative surgery.

This peculiar property of variation in gray levels of liver and liver lesions can be used to isolate liver lesions from the liver. Isolating abdominal organs like liver from other organs in the abdomen or isolating liver lesion from the liver is accomplished using image segmentation. Segmentation is performed on the basis of gray level similarity or dissimilarity [2]. Based on dissimilarity point, line and edge, are detected but these features are not definite in abdominal CT images and difficult to locate. The lesions usually are closed curves and having a definite gray level range which is different from the liver as a whole, so regional segmentation methods based on similarity of gray levels are more suitable. Among these, region growing or region split and merge involve accurate selection of seed point, which is difficult to obtain. Hence in the proposed method thresholding technique is used to segment lesions present in liver. The size of liver, the variations in gray level between liver and liver lesion and type of lesion may vary from person to person. Thus, with threshold based segmentation, choice of appropriate thresholds for each case becomes a crucial task. Also, presence of similar gray levels in adjoining liver regions makes liver lesion segmentation further difficult and challenging.

Precise segmentation of liver lesions enhances manual analysis of radiographic images to assess impact and extent of lesion on neighbouring regions. Once the lesions become accessible, their invariant features like area, centroid etc can be obtained and used for further processing like image registration, image fusion etc. In this paper an automatic threshold selection technique based on statistical measures of CT images, is proposed to segment liver lesions in the pre contrast and post contrast image slices of the patient. The proposed automatic thresholding technique is tested on various types of lesions and segmented lesions accurately.

Literature related to liver segmentation is discussed in section 2, proposed automatic threshold based segmentation technique is explained in section 3, performance of proposed algorithm for different types of lesion is evaluated in section 4 and section 5 gives conclusion.

2. RELATED WORK

Organ segmentation in abdominal CT images is challenging due to beam hardening artifacts seen as focal areas of low attenuation near bones, partial-volume artifacts giving blurred images, streak artifacts due to patient movement, respiration etc. Lav R. Varshney [3] presented survey of various abdominal organ segmentation methods for CT scan images based on mathematical morphology, thresholding, neural network, level set methods, model fitting and data directed methods. Lav R. Varshney summarised limitations of different abdominal organ segmentation methods. Model based methods normally use rib cage as reference to obtain location of organs. Rule based methods make use of organ invariants. Morphological operations segment abdominal CT images well and are usually applied in combination with other methods [3]. Hepatic tissues from abdominal CT images have been classified using various texture measures [4]. In this method radiologist specifies the Region of Interest (ROI) (hepatic tissues) located in non enhanced images. Statistical measures are used to determine texture of ROI. Feature vectors of these ROI are formed, which are reduced using Genetic Algorithm. These feature vectors using five different kinds of neural networks, classifies ROI into four types of tissues. Neural Network (NN) based methods using texture of organs as features become ineffective in organ segmentation because organs themselves have substructures having different structures and texture may not be properly defined [4]. Also, neural network based methods need lot of training, which increases the complexity of the technique. Thresholding based methods have limited utility since various organs and tissues may have similar gray levels. Despite this, segmentation of contrast enhanced CT images, defined by a body box [5] strictly enclosing abdominal structures has improved the results. In [5] significant edges are determined in the three coronal, sagittal and axial slices to segment the heart volume so that information of only liver can be obtained from the images. This was required because the slices involved presence of heart and liver in them. Liver gray levels are estimated from histograms of the images. Major peak of histogram identifies liver region. First minima on either side of this peak give the minimum and maximum gray levels to

segment the liver. This method is unable to isolate lesions from the liver. Moreover to segment liver lesions, a different set of thresholds are required as they have different gray levels compared to those of liver. In [6] a local multiphase C-V level set algorithm is integrated with multilevel Otsu's method to segment complicated liver lesion. Here points close to lesion are identified manually on the image to construct a local liver lesion image. Then lesion is pre segmented with multilevel Otsu's method. It gives the initial contour of lesion. Final contour of lesion is then obtained using multiphase C-V level set method. The method categorises the lesion into normal liver tissue, the enhanced part of lesion and the unenhanced part of lesion but formation of initial image of lesion requires points to be selected manually.

Soler et al [7], presented fully automatic anatomical, pathological and functional segmentation of liver from spiral CT scan images for hepatic liver surgery. In [7] all neighbouring organs like spleen, aorta, skin, bones etc are first delineated by use of information of histograms of specific regions. A model of the liver is then merged with that of segmented liver. Reconstructed liver provided information useful for pre operative surgery. This method also, isolates the liver from other organs but it is unable to isolate lesions within liver.

Region growing techniques are useful because they club pixels with similar gray levels but it is difficult to choose the seed pixel. In [8] automatic liver segmentation of contrast enhanced CT images is carried out. It uses an advanced form of region growing method and is able to segment liver parenchyma in most of the cases except where liver lesions are very large. Another texture based volumetric region growing method for liver segmentation in [9] finds seed point automatically and also the threshold for stopping the region growing process. In the pre processed 3D Abdominal CT images, seed point is chosen as the centre of CT volume and ROI is determined as abdominal pixels with orientation ranging between 90-120 degree from coordinates of abdominal centre of mass using a cost minimization approach. Local texture of neighbouring volume of seed pixel is determined from texture features, which are calculated from gray level co-occurrence matrix. Euclidean distances determine homogeneity between voxels with 6-connectivity. Every 6-neighbour voxel with euclidean distance below a threshold is included in the volume. Stopping threshold is a function of the Gaussian parameters as distance from seed to voxels is a Gaussian distribution. The method has been applied to both basal and arterial images but not for the same slice of the two images. This method gives under segmentation near liver boundaries.

Snake based liver lesion segmentation is useful for MR images [10]. In the MR images speckle is removed through median filtering and snake of liver contour is found by use of edges. Region is grown by the seed which is lesion pixel. Lesion boundary information is used for the snake model. Fuzzy edge information distinguishing lesion and liver region is used to modify the snake recursively to obtain accurate lesion boundaries. This method proves insufficient for high texture and low contrast content in images. In [11] watershed and active contour algorithms have been applied for segmentation to obtain volumetric information of hepatic liver metastases in CT images. Comparison of these algorithms showed that active contour algorithm is able to detect a good number of lesions compared to watershed algorithm. Segmentation errors occurred in case of lesions lying towards liver periphery. Watershed algorithm causes over segmentation. Over segmentation problem resulting due to use of watershed segmentation was considerably reduced in [12]. In this method watershed segmentation is applied to liver cancer lesion segmentation. Initially image gradient is found using sobel operator. Thresholding was done using Otsu's method followed by watershed segmentation to reduce its complexity. The segmented regions obtained are then merged by use of region adjacency graphs to merge most similar regions. Average gray scale and boundary average guideline are used as parameters to match the segmented regions. Though over segmentation problem of watershed segmentation is resolved to some extent, still threshold selection for region merging is dependent on manual expertise. In all the segmentation methods discussed, segmentation of liver lesions is performed in very few cases and that too for a single type of lesion. Moreover both pre contrast and post contrast images are not considered for segmentation. In the next section an automatic thresholding technique based on statistical measures is proposed to segment different types of liver lesions in

both pre contrast and post contrast images. This segmentation is useful in determining feature points lying on the lesion boundary and within the lesion.

3. AUTOMATIC THRESHOLD BASED SEGMENTATION

The proposed method uses three steps Automatic thresholding, Morphological operations and Boundary extraction for segmentation in abdominal 2D-CT pre contrast and post contrast images. A flow diagram of the proposed method is as shown in Figure 1.

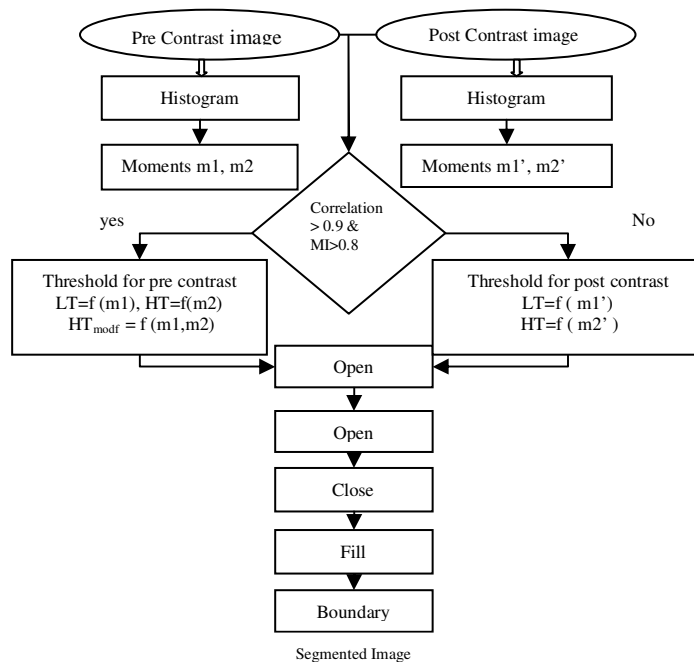


FIGURE1: Flow diagram of Automatic threshold based segmentation

3.1 Automatic Thresholding

It is observed that selection of a single threshold for segmentation of lesions in liver is not sufficient. The intensity of the entire liver region varies with respect to that of the liver lesions and type of lesion. A minimum of two thresholds are needed to segment liver and lesions from the abdominal CT images. A single threshold segments the liver as a whole and is incapable to segment lesions present in the liver. Paola et al [5] identified the patient body by using the minimum to the right of first gray level peak of the histogram of original image as threshold. Then body box defining the liver was obtained. The highest maxima of the histogram of this body box represent the liver and the first local minima lying on either side of the maxima represent gray levels corresponding to the liver. Paola et al [5] segmented liver but the method cannot segment lesions from liver. In this paper, an automatic threshold selection method is proposed to segment various types of lesions that can be present in a liver.

In the proposed method the lesions within the liver are determined using statistical measures. In the proposed algorithm the two thresholds used to segment lesions are called Lower Threshold (LT) and Higher Threshold (HT). The mean value of the image is used to determine LT and another statistical measure i.e. standard deviation is used to determine HT to segment the liver and lesion within the liver. These thresholds are minima lying on either side of the single gray level peak (P_G) of the histogram of the pre and post contrast slices for different patients with different types of lesions such as metastases, ascitis and multiple liver metastases and cyst. Precise calculation of these thresholds is crucial because they define the accuracy of segmentation of lesions, which in turn affect the accuracy of the results of further processing on segmented images. It is observed that, in most of the sets of pre contrast and post contrast

images, number of maxima in pre contrast and post contrast images are equal and for such cases the HT lies to the right of P_G of the histogram, while the LT usually lies to the other side of P_G i.e. left of P_G . However exceptionally, when in the pre contrast images, the difference in gray level between liver and lesion is barely distinguishable and there is large difference in mean value of the pre contrast and the post contrast image slices, in such cases the LT and HT for post contrast image, both lie to the same side of P_G . In such cases the total number of maxima in the pre contrast and post contrast images is unequal and both the LT and HT lie to the right of the P_G . For these exceptional cases, the LT for pre contrast and post contrast image is kept same but the HT for pre contrast image needs modification.

In the proposed method, LT and HT are obtained from the statistical properties of the intensity histogram using statistical moments. Statistical moments are measures which provide the texture information [13]. The shape of histogram can be described by use of central moments μ_n defined in Eq (1)

$$\mu_n = \sum_{i=0}^{L-1} (z_i - m)^n p(z_i) \quad (1)$$

In Eq (1) L is the number of possible gray levels, z_i is the discrete random variable indicating intensity levels in an image, n is the order of the moment and $p(z_i)$ is the histogram of the intensity levels in the region i.e. estimate of the probability of occurrence of intensity value z_i . LT is defined as twice the measure of average intensity of the histogram i.e. mean (m).

$$LT = 2m \quad (2)$$

where mean (m) is expressed as

$$m = \sum_{i=0}^{L-1} z_i p(z_i) \quad (3)$$

HT is defined as twice of the measure of average contrast i.e. standard deviation (σ),

$$HT = 2\sigma, \quad (4)$$

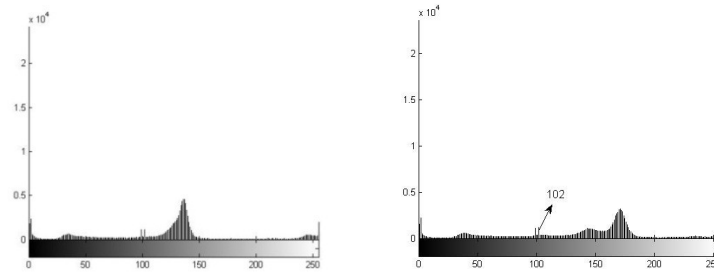
standard deviation σ is given as in Eq(5)

$$\sigma = \sqrt{\mu_2(z)} = \sqrt{\sigma^2} \quad (5)$$

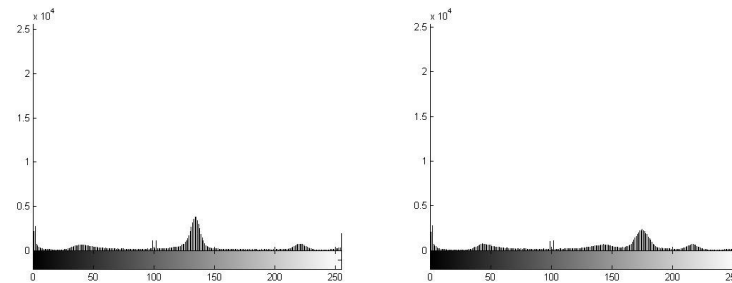
In exceptional cases for which number of peaks in the histograms of pre and post contrast image slices of patients are unequal, HT for pre contrast image is modified to HT_{modf} as given in Eq(6).

$$HT_{modf} = 2LT - \frac{(HT - LT)}{2} \quad (6)$$

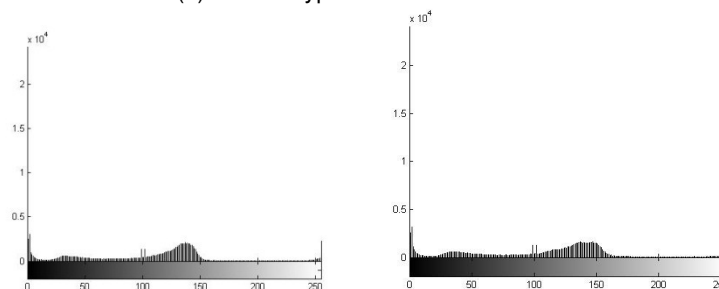
This can be explained through the histograms for the selected slices of few cases (patients) suffering from ascitis, liver metastases and multiple liver metastases and cyst. Histograms of chosen slices of these patients are shown in Figure 2.



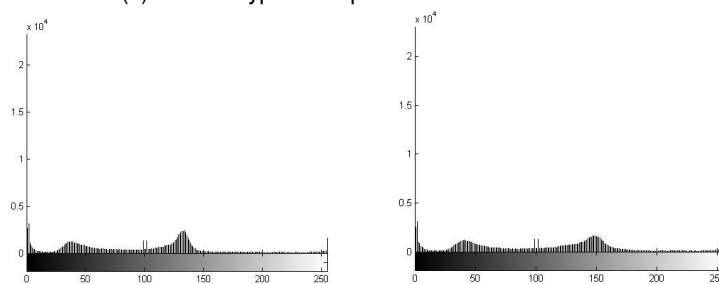
2(a) Lesion Type: Ascitis



2(b) Lesion Type: Liver Metastases



2(c) Lesion Type: Multiple Liver metastases



2(d) Lesion Type: Cyst

FIGURE2: Histograms of Pre contrast images (left column) and Post contrast images (right column) for different types of lesions in patients p1-p4. 2(a) ascitis: patient p1, 2(b) liver metastases: patient p2; 2(c) multiple liver metastases: patient p3; 2(d) cyst; patient p 4

In Figure 2, histograms in the left column are of pre contrast image slices and histograms in the right column are of post contrast image slices of the patients p1-p4. In the pre and post contrast images of most cases, the thresholds LT and HT calculated according to proposed method are seen to lie on either side of the single peak (P_G) in the histogram appearing at a gray level of 102 i.e. LT lies to the left of P_G and HT lies to the right of P_G . However, as identified earlier, there are some exceptional cases, when LT and HT both lie on the same side of P_G . This can be seen in histograms of Figure 2(a) and Figure 2(b). Here both LT and HT for post contrast image, lie to the right of P_G (102) because of the large difference in mean value of the pre contrast and the

post contrast image slices. In the slices of histograms of pre contrast images in Figure 2(a) and Figure 2(b) the gray level differences between liver and lesion is barely distinguishable whereas for slices of histograms of pre contrast images in Figure 2(c) and Figure 2(d), the gray level difference in liver and lesion is considerable. Thus the threshold levels for pre contrast images of Figure 2(a) and Figure 2(b) is different from that for Figure 2(c) and Figure 2(d). All these commonly indicate feeble difference in gray level of liver and lesion. It is also observed that the number of maxima in the histograms of the pre and post contrast images in Figure 2(c) and Figure 2(d) are equal whereas in Figure 2(a) and 2(b) are unequal. This is attributed to the availability of relatively less gray levels in pre contrast image and more gray levels in post contrast images for patient p1 and patient p2. In addition a higher correlation and higher Mutual information between the pre and post contrast images is seen (above 0.9 and above 0.8 respectively) along with large value of mean for the images for p1 and p2, as shown in Table 1. In case of patient p3 and patient p4, the number of gray levels in pre and post contrast images being relatively same, and histogram peaks are also equal in number.

Patient	Correlation	Mutual Information	Mean (pre contrast)	Mean(post contrast)
p1	0.9430	0.8839	50.0388	58.9335
p2	0.9106	0.8451	45.1442	51.9194
p3	0.6769	0.7581	43.4805	45.4769
p4	0.8761	0.8747	41.5193	45.9020

TABLE 1: Correlation coefficient and Mutual Information of pre and post contrast images for patients p1-p4

Hence modification in HT according to Eq (6) has been suggested for those pre contrast images for which correlation coefficient is above 0.9, Mutual Information above 0.8 and it has a higher mean value. In pre contrast images i.e. Figure 2(c), 2(d); HT is calculated according to Eq (5). In all post contrast images LT is calculated according to Eq (3) and HT is calculated according to Eq (5) only. The Lower threshold (LT) and Higher threshold (HT) calculated using proposed method for each patient are as given in Table 2.

Patient	Pre Contrast			Post Contrast	
	Lower Threshold (LT)	Higher Threshold (HT)	Modified Higher Threshold (HT _{modf})	Lower Threshold (LT)	Upper Threshold (HT)
p1	100.0776	141.3447	131.0279	117.8670	157.9767
p2	90.2883	140.5499	127.9845	103.8389	156.481
p3	86.9609	127.9678	Not Applicable	90.9538	133.1248
p4	83.0385	122.5669	Not Applicable	91.8039	132.2971

TABLE 2: LT and HT for pre and post contrast images

Using this proposed automatic thresholding for 2D abdominal CT images it is seen that apart from segmented lesions, those regions of the abdomen that have similar gray levels as the gray levels of lesions are also segmented. The region of interest however is the liver. So the geometrical location of liver is considered and inter patient variations in size of the liver are taken into account to determine the region of interest. The regions other than liver that are segmented are masked by choosing a definite boundary box which does not include the entire abdomen as in [5] but encloses only the region to the left of aorta containing the liver as shown in Figure 3. Usually region between the coordinates (100:300) and (100:250) correctly encompass the right lobe of the liver.

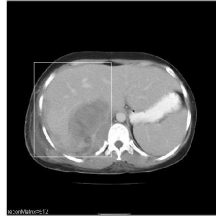


FIGURE 3: Boundary box enclosing the liver

The image thus obtained after masking now contains only the lesions in the liver mainly with a little segmentation seen beyond the rib cage. After thresholding some morphological operators are applied.

3.2 Morphological Processing

Morphological operations involve two rounds of 'opening' followed by 'closing', then a 'fill' operation followed by 'majority' operation. Morphological opening completely removes regions of an object that cannot contain the structuring element, smoothes inside of object contours, breaks thin connections and removes thin protrusions [3]. This elimination of certain connections removes some small segmented regions which are not the lesions. It filters details and simplifies image by rounding corners. Morphological closing tends to smooth contours of objects. Unlike opening however, it joins narrow breaks, fills long thin gulfs and fills holes smaller than the structuring element. The 'fill' operation fills the isolated individual pixels. 'Majority' operation sets a pixel to 1 if 5 or more pixels in its 3x3 neighbourhood are 1 else it sets the pixel to 0. After application of morphological operators, the lesions are then identified by precise boundaries obtained using boundary detection.

3.3 Boundary Detection

The boundaries of lesions after Morphological Processing are determined using 'bwboundaries' function in MATLAB. This function traces the exterior boundaries of objects, as well as boundaries of holes inside these objects, in the binary image. Liver lesions are of different types with different shapes and size. Hence two thresholds are needed to define them distinctly. From the set of images under consideration, it is observed that in order to identify liver lesions having very small size as in case of liver metastases and multiple liver metastases, in the proposed method boundary lengths greater than 2 of lesion are considered as threshold. However for lesions with large sizes as in case of ascitis, enlarged liver etc. boundary lengths greater than 200 are used as threshold for segmentation. These thresholds for boundary length are found by experiments carried out over a large number of images.

4. RESULTS

The proposed method is tested on 20 image slices of 2D abdominal CT images having varied liver lesions like liver metastases, ascitis, cyst and multiple liver metastases. The single slice spiral CT, axial images are taken from WIPRO GE CT/e machine. The images are in DICOM format and exported to .JPG format using e-Film 1.5.3 workstation. These images are converted to gray scale images to provide a set of 2D slices of size 512x512 with slice thickness 10mm. Number of slices per patient varied from 20 to 58. The experiment was performed on 32 bit, 2GHz. Intel Core 2 Duo processor using MATLAB. In each patient, corresponding slice of pre contrast and post contrast image is selected to segment the liver lesions. No pre processing is required to implement the proposed algorithm. The LT and HT for pre contrast and post contrast images in each patient are determined according to the proposed method. Boundaries of segmented lesions in both pre contrast and post contrast images of all patients are depicted as in Figure 4 -7

Analysis of segmented lesions of pre and post contrast images of the four peculiar cases (patients) is done to evaluate the performance of the proposed algorithm, of which pre contrast images of 2 patients need modification in HT and pre contrast images of 2 patients do not need modification in HT. In Figure 4 and Figure 5, pre contrast images p1(a) and p2(a) of patient 1 and patient 2 having ascitis and liver metastases are shown. These images are segmented using LT and HT, the segmented regions so obtained are shown in blue in p1(b) and p2(b), which are segmenting entire liver region and unable to show segmented lesions. However when these images are segmented using LT and HT_{modf} according to the Eq(6) and as per the requirement for modification in threshold suggested in pre contrast images when correlation coefficient of pre and post contrast images is greater than 0.9 and mutual information (MI) above 0.8, then the lesions are clearly segmented in these pre contrast images which are shown in p1(c) and p2(c). The segmented regions in image p1(b) and p1(c), p2(b) and p2(c) clearly indicate the requirement of modification in HT of these pre contrast images. The post contrast images of the same patients are shown in p1(d) and p2(d). Segmented lesions on these post contrast images using LT and HT as per the proposed algorithm are shown in red as in p1(e) and p2(e)

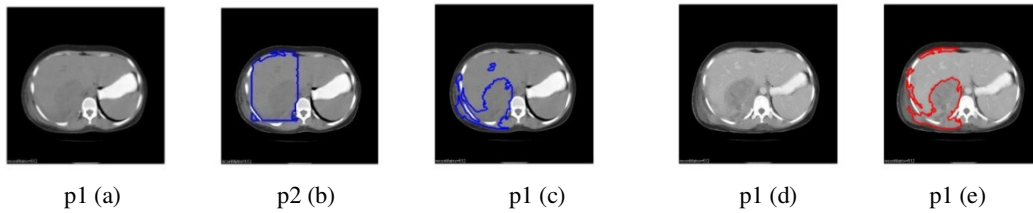


FIGURE 4: Details of patient p1: (a) Pre contrast image (b) Boundaries of lesions segmented with LT, HT in pre contrast images (c) Boundaries of lesions segmented with LT, HT_{modf} in pre contrast images (d) Post contrast images (e) Boundaries of lesions segmented with LT, HT in post contrast images

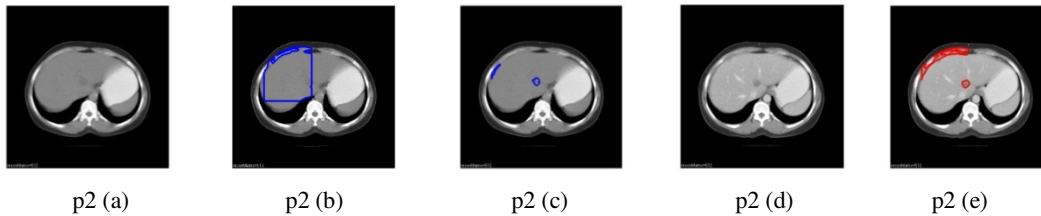


FIGURE 5: Details of patient p2: (a) Pre contrast image (b) Boundaries of lesions segmented with LT, HT in pre contrast images (c) Boundaries of lesions segmented with LT, HT_{modf} in pre contrast images (d) Post contrast images (e) Boundaries of lesions segmented with LT, HT in post contrast images

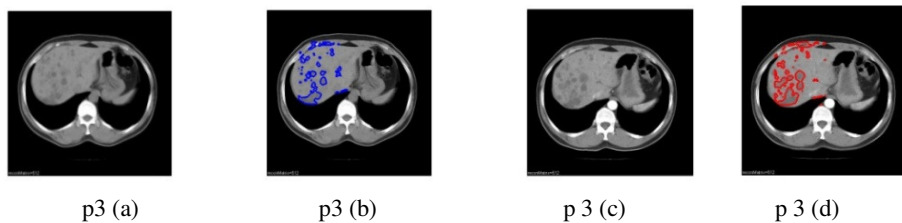


FIGURE 6: Details patient p3 (a) Pre contrast images (b) Boundaries of lesions segmented with LT, HT in pre contrast images (c) Post contrast images (d) Boundaries of lesion segmented with LT, HT in post contrast images

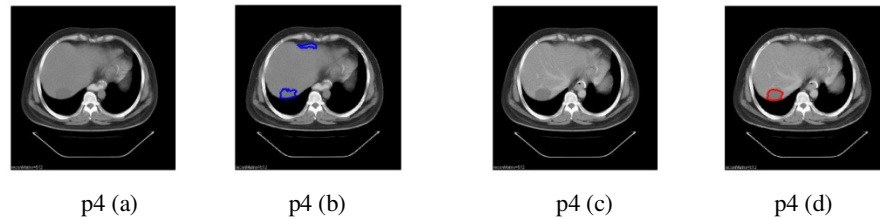


FIGURE 7: Details of patient p4. (a) Pre contrast images (b) Boundaries of lesions segmented with LT, HT in pre contrast images (c) Post contrast images (d) Boundaries of lesion segmented with LT, HT in post contrast images

In figure 6 and Figure 7 pre contrast images of patients p3 and p4 with multiple liver metastases and cyst are shown in p3 (a) and p4 (a) respectively. These pre contrast images are segmented using LT and HT and the segmented lesion boundaries on the respective pre contrast images are shown in blue in p3 (b) and p4 (b). The post contrast images for patients p3 and p4 are shown in p3 (c) and p4 (c). These post contrast images are segmented using LT and HT and segmented lesion boundaries are shown in red on respective post contrast images in p3 (d) and p4 (d). These observations clearly show that within the defined boundary box, similar lesions in both pre contrast and post contrast images can be readily obtained by the proposed segmentation method. The lesions detected by the proposed method of segmentation can assist in precise determination of feature points. Feature points like Hull points of segmented regions and their centroids which are invariant to translation and rotation of segmented post contrast image. For illustration these have been shown only for patient p2 are as shown in Figure 6.

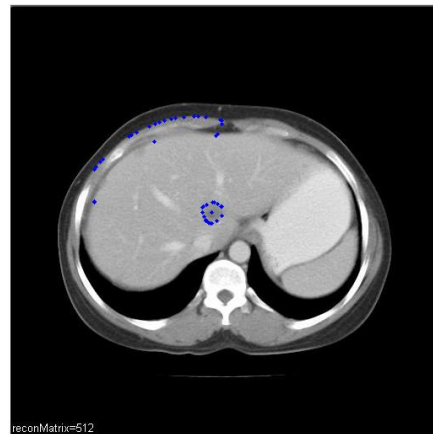


FIGURE 6: Hull points and centroid of segmented lesion

5. CONCLUSIONS

An automatic threshold based proposed method is able to segment pre contrast and post contrast abdominal 2D CT images having varied types of lesions like liver metastases, ascitis and cyst more accurately. LT and HT are used to segment the lesions from the liver. In some cases, pre contrast images have very feeble difference in gray levels of liver and liver lesions but for some cases pre contrast images show relatively larger gray level difference. Proposed HT modification in pre contrast images with feeble gray level difference in liver and liver lesion enhances visual assessment of liver lesion in pre contrast images. The objective of the paper to locate lesion boundaries to find feature points on and within the lesion for further processing is accurately accomplished. The lesions are commonly identified in both the pre contrast and post contrast images in each patient. A boundary box is used to restrict the region of interest. This boundary box can be used in future for segmenting lesions of surrounding anatomy.

Acknowledgement

We sincerely acknowledge the efforts of Dr. P.K Dave, MD, DRM, Radiologist, Kamla Nehru Gas Relief Hospital Bhopal, India, for his support and authorities of Jawaharlal Nehru Cancer Hospital and Research Center Bhopal, India, for providing the images for the case study.

REFERENCES

- [1] L. Sandra, Hagen-Ansert, William J. Zwiebel. "Textbook of Diagnostic Ultrasonography", Elsevier Health Sciences, pp. 127(2006)
- [2] S. Annadurai, R. Shanmugalakshmi, "Fundamentals of Digital Image Processing", Pearson Education, pp. 228, First impression (2007).
- [3] Lav R. Varshney. "Abdominal Organ Segmentation in CT scan Images: A Survey". pp. 1-4, Cornell University, August 2002.
- [4] S. Gr. Mougiakakou, I.Valavanis, K. S. Nikita, A. Nikita, and D. Kelekis. "Characterization of CT Liver Lesions Based on Texture Features and a Multiple Neural Network Classification Scheme". In Proceedings of the 25th Annual IEEE Int. Conf. Engineering & Medicine Biology Society, pp. 1287-1290, September 2003
- [5] P. Campadelli, E. Casiraghi, and G. Lombardi, "Automatic liver segmentation from abdominal CT scans". In Proceedings of the 14th IEEE International Conference on Image Analysis and Processing, Modena, pp. 731-736, September 2007
- [6] Zhaohui Luo, Xiaoming Wu, Renjing Cen, Shanxing Ou. "Segmentation of Complicated Liver Lesion Based on Local Multiphase Level Set". In Proceedings of the 3rd IEEE International Conference on Bioinformatics and Biomedical Engineering, Beijing, pp.1-4, 2009
- [7] L Soler, H. Delingette, G. Malandain, J. Montagnat, N. Ayache, C. Koehl, O. Dourthe, B. Malassagne, M. Smith, D. Mutter, and J. Marescaux. "Fully automatic anatomical, pathological, and functional segmentation from CT scans for hepatic surgery". *Computer Aided Surgery*, 6(3): pp. 131-142, 2001
- [8] L. Ruskó, G. Bekes, G. Németh, and M. Fidrich. "Fully automatic liver segmentation for contrast-enhanced CT images". T. Heimann, M. Styner, B. van Ginneken (Eds.): 3D Segmentation in The Clinic: A Grand Challenge, Hungary, pp. 143-150, 2007
- [9] O. Gambino, S. Vitabile, G. Lo. Re, G. La. Tona, S. Librizzi, R. Pirrone, E. Ardizzone, and M. Midiri. "Automatic Volumetric Liver Segmentation Using Texture Based Region Growing". In Proceedings of the International Conference on Intelligent and Software Intensive Systems, Krakow, pp. 146-152, 2010
- [10] C. Krishnamurthy, J.J. Rodriguez, and R.J. Gillies. "Snake-Based Liver Lesion Segmentation". In Proceedings of the 6th IEEE South West Symposium on Image Analysis and Interpretation, Lake Tahoe, CA, pp. 187-191, 2004
- [11] P.J. Yim and D.J. Fora. "Volumetry of Hepatic Metastases in Computed Tomography using the Watershed and Active Contour Algorithms". In Proceedings of the 16th IEEE Symposium on Computer-Based Medical Systems, New York, pp.329-335, 2003
- [12] J. Liu, Z. Wang, R. Zhang. "Liver Cancer CT Image Segmentation Methods based on Watershed Algorithm". In Proceedings of the International Conference on Computational Intelligence and Software Engineering, Wuhan, pp. 1-4, 2009
- [13] R.C. Gonzalez, R. E. Woods, and S.L Eddins. "Digital Image Processing Using MATLAB", Pearson Education, pp. 167-170, 478(2007)

Segmentation Based Multilevel Wide Band Compression for SAR Images Using Coiflet Wavelet

Dr.P.Subashini

Associate Professor

Department of Computer Science

Avinashilingam Deemed University for Women

Coimbatore, 641 043, India

mail.p.subashini@gmail.com

M.Krishnaveni

Research Assistant

Department of Computer Science

Avinashilingam Deemed University for Women

Coimbatore, 641 043, India

krishnaveni.rd@gmail.com

Abstract

Synthetic aperture radar (SAR) data represents a significant resource of information for a large variety of researchers. Thus, there is a strong interest in developing data encoding and decoding algorithms which can obtain higher compression ratios while keeping image quality to an acceptable level. In this work, results of different wavelet-based image compression and segmentation based wavelet image compression are assessed through controlled experiments on synthetic SAR images. The effects of dissimilar wavelet functions, number of decompositions are examined in order to find optimal family for SAR images. The choice of optimal wavelets in segmentation based wavelet image compression is coiflet for low frequency and high frequency component. The results presented here is a good reference for SAR application developers to choose the wavelet families and also it concludes that wavelets transform is rapid, robust and reliable tool for SAR image compression. Numerical results confirm the potency of this approach

Keywords: Image Processing, Compression, Segmentation, Wavelets, Quality Measures, SAR Images

1. INTRODUCTION

The necessity of efficient and consistent compression techniques for remote sensing imagery is increasing rapidly on the bases of network transmission [2]. In image processing domain the neglectation of peculiar characteristics of the data accounts a lot in poor recognition system [9]. The usual approach to deal with this problem is to resort to edge-preserving filters. Segmentation, compression and filtering should be carried out jointly for extraction and better representation of most relevant features. Therefore a fundamental approach to image filtering and compression requires the prior segmentation of the image in homogeneous regions [5]. This work includes the strong background of research result proposed by Zhaohui Zeng and Ian Cumming for SAR image compression using DWT [15]. Here wavelet-based despeckling is used implicitly to distinguish between noise components (high and low) and its state boundaries[4][2]. This work aims at studying and quantifying the potential advantages provided by wavelet towards segmentation based image compression. This work is investigated from few SAR image collections like ERS Synthetic Aperture Radar (SAR) imagery , ESA,JERS and radarsat . The paper follows as. Section 2 explains the compression schemes needed for SAR segmentation. Section 3 deals with wavelet analysis and its metrics which comprises the performance measures of wavelet families. Section 4 describes the compression coding schemes based on segmentation. Section 5 presents and discusses the results of a number of experiments, and Section 6 draws conclusions with possible scenario.

2. SAR IMAGE WAVELET COMPRESSION

In the scientific literature concerning SAR system, the term compression is often used to indicate the data processing which allows focusing the received echoes. SAR domain compression[3][14] is carried out off-line at the base station, which can leverage on more time and computation power. So image compression became an operating instrument, for scientific research[9]. In basic literature evaluation is done on quality metrics, neglecting the visual effects and the consequences on the image interpretability. In order to attain above conditions and limitations of SAR, new compression techniques on wavelet transform is characterized with good time-frequency localization, multiresolution representation low-complexity implementation through filter banks [5]. Since the wavelet transform is a bandpass filter with a known response function (the wavelet function), it is possible to *reconstruct* the original time series using either deconvolution or the inverse filter [15]. This is straightforward for the orthogonal wavelet transform (which has an orthogonal basis), but for the continuous wavelet transform it is complicated by the redundancy in time and scale. However, this redundancy also makes it possible to reconstruct the time series using a completely different wavelet function, the easiest of which is a delta (δ) function[1]. In this case, the reconstructed time series is just the sum of the real part of the wavelet transform over all scales which is stated in eqn(1)

$$x_n = \frac{\delta_j \delta_t^{1/2}}{C_\delta \psi_0(O)} \sum_{j=0}^J \frac{\Re\{W_n(s_j)\}}{s_j^{1/2}} \quad \text{----(1)}$$

The factor $\psi_n = (0)$ removes the energy scaling, while the $s_j^{1/2}$ converts the wavelet transform to an energy density. The factor C_δ comes from the reconstruction of a δ function from its wavelet transform using the function $\psi_n = (n)$. This C_δ is a constant for each wavelet function .Note that if the original time series was complex, then the sum of the complex $W_n(s)$ would be used instead [3][5]. To derive C_δ for a new wavelet function, first assume a time series with a δ function at time $n = 0$, given by $x_n = \delta_n$. This time series has a Fourier transform $s\omega_k = N-1$, constant for all k . Substituting $s\omega_k$ into (1), at time $n = 0$ (the peak), the wavelet transform becomes

$$W_\delta(S) = \frac{1}{N} \sum_{k=0}^{N-1} \psi^*(s\omega_k) \text{-----} (2)$$

The reconstruction eqn (2) then gives eqn (1) .The C_δ in eqn (1) is scale independent and is a constant for each wavelet function.

$$\sigma^2 = \frac{\delta_j \delta_t}{C_\delta N} \sum_{x=0}^{N-1} \sum_{j=0}^J \frac{|W_n(s_j)|^2}{s_j} \text{-----(3)}$$

The total energy is conserved under the wavelet transform, and the equivalent of *Parseval's theorem* for wavelet analysis is given in eqn(3).Both eqn (1) and eqn (3) should be used to check wavelet routines for accuracy and to ensure that sufficiently small values of s_j and δ_j have been chosen.

3. WAVELET ANALYSIS

This section describes the method of wavelet analysis, includes a discussion of different wavelet functions, and gives details for the analysis of the wavelet power spectrum. Results in this section are adapted to Segmentation based wavelet compression from the continuous formulas given in coiflets. The wavelet bases performed over SAR images are as follows. Haar wavelet is the simplest of the wavelet transforms[10]. This transform cross-multiplies a function against the Haar wavelet with various shifts and stretches, like the Fourier transform cross-multiplies a function against a sine wave with two phases and many stretches. Daubechies wavelets are a family of orthogonal wavelets defining a discrete wavelet transform and characterized by a maximal number of vanishing moments for some given support. Meyer's wavelet construction is

fundamentally a solvent method for solving the two-scale equation. Symlet wavelet is only nearly symmetric, and is not exactly symmetrical. Coiflets are discrete wavelets designed by Ingrid Daubechies, to have scaling functions with vanishing moments. Biorthogonal wavelet is a wavelet where the associated wavelet transform is invertible but not necessarily orthogonal. Reverse biorthogonal is a spline wavelet filters.

To verify the validity of the wavelet families the results are compared based on PSNR ratio and MSE. With the extension of the work, subjective evaluation is also done based on the decomposition level. Compression is been the main phenomena for the evaluation job [8][12]. The DWT was used with coiflets, least asymmetric compactly-supported wavelet with eight vanishing moments with four scales. The 120 x 120 pixel region SAR images are used for applying wavelet families. They were contaminated with speckle noise[2][3].

Some of the parameters taken for analysis of wavelet on SAR images are Mean Square Error, Peak Signal to Noise Ratio, Compression Score and Recovery Performance.

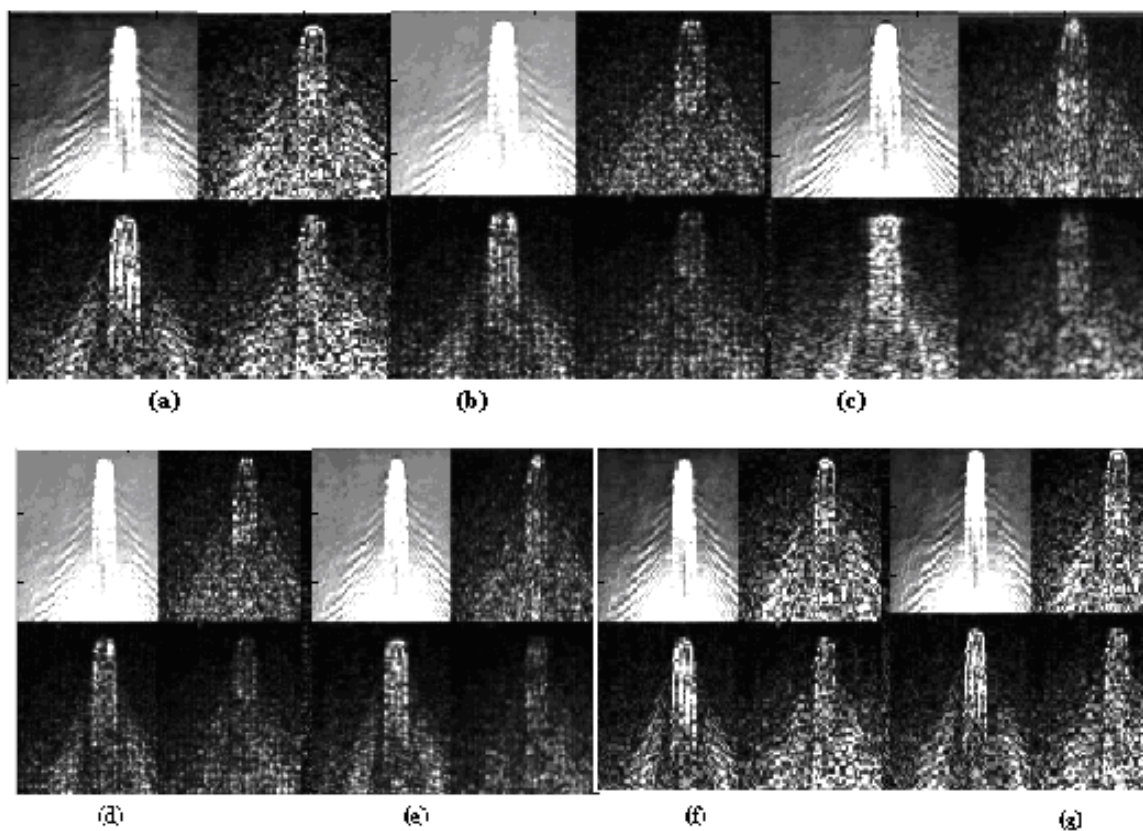


FIGURE 1: (a) Haar wavelet (b) Daubechies wavelet (c) Meyer Wavelet (d) Symlet wavelet (e) Coiflets wavelet (f) Bi orthogonal wavelet (g) Reverse Bi orthogonal

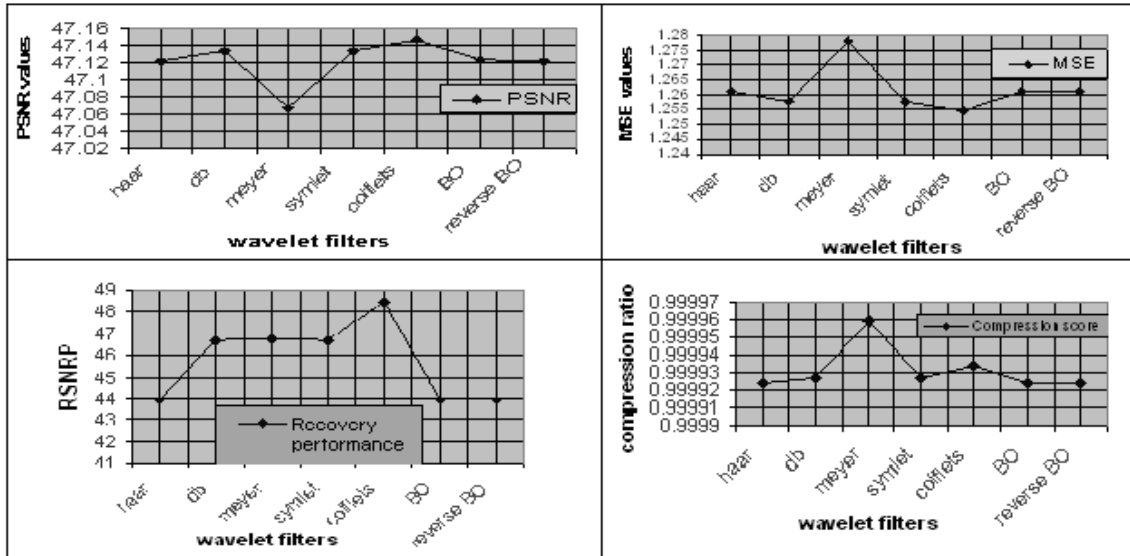


FIGURE 2: (a) Peak Signal to Noise Ratio (b) Mean Square Error Rate (c) Recovery Performance (d) Compression Score

A sample of one SAR image is subjectively explained by the results in the above figure (1) and objectively in figure (2). From the above got results the coiflets of wavelet families out performs the rest of the wavelet families.

4. SEGMENTATION BASED WAVELET COMPRESSION

The idea behind wavelet image compression, like that of other transform compression techniques is, it removes some of the coefficient data from the transformed image [8]. Encoding may be applied to the remaining coefficients. The compressed image is reconstructed by decoding the coefficients, if necessary, and applying the inverse transform to the result[17]. No much image information is lost in the process of removing transform coefficient data.

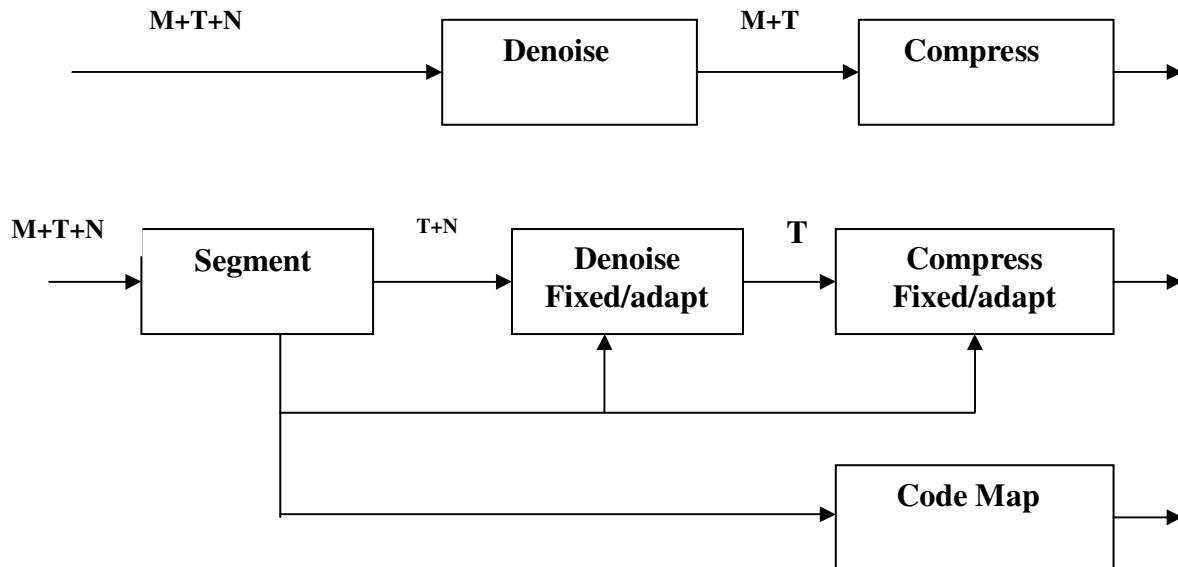


FIGURE 3: Segmentation based encoding schemes

This process of image compression through transformation and information removal provides a useful framework for understanding wavelet analysis[8]. Mathematical references develop

multiresolution analysis through the definition of continuous scaling and wavelet functions[18]. In order to compare the performance of wavelet and contourlet transform in our proposed compression scheme, we compute the compression ratios (CR) for various quantization levels. CR is defined as the ratio between the uncompressed size and the compressed size of an image [13]. To compute the CR in a fairly way, the original image is encoded using the resulting number of bits it saved. This compression eliminates almost half the coefficients, yet no detectable deterioration of the image appears. The compression procedure contains three steps: (i) Decomposition. (ii) Detail coefficient thresholding[16] ,for each level from 1 to N, a threshold is selected and hard thresholding is applied to the detail coefficients, and(iii) Reconstruction.

5. PERFORMANCE EVALUATION

Fig 4 and 5 shows PSNR and compression ratio curves for various SAR images[3] through compression techniques. Fig 6 and 7 shows the Normalized cross correlation and Normal Absolute error rate for the sample SAR images between two methods wavelet based compression (WBC) and segmentation based wavelet compression (SBWC). The image size held is 256x256. These curves plot compression ratio versus image quality, as represented by PSNR. The most desirable state on this graph is the upper right quadrant, where high compression ratios and good image quality live.

From the Figure 6 and 7 the Cross-correlations help identify variables which are leading indicators of other variables or how much one variable is predicted to change in relation the other variable. The identified method is justified in all the above taken parameters which objectively states the phenomena in prominent manner

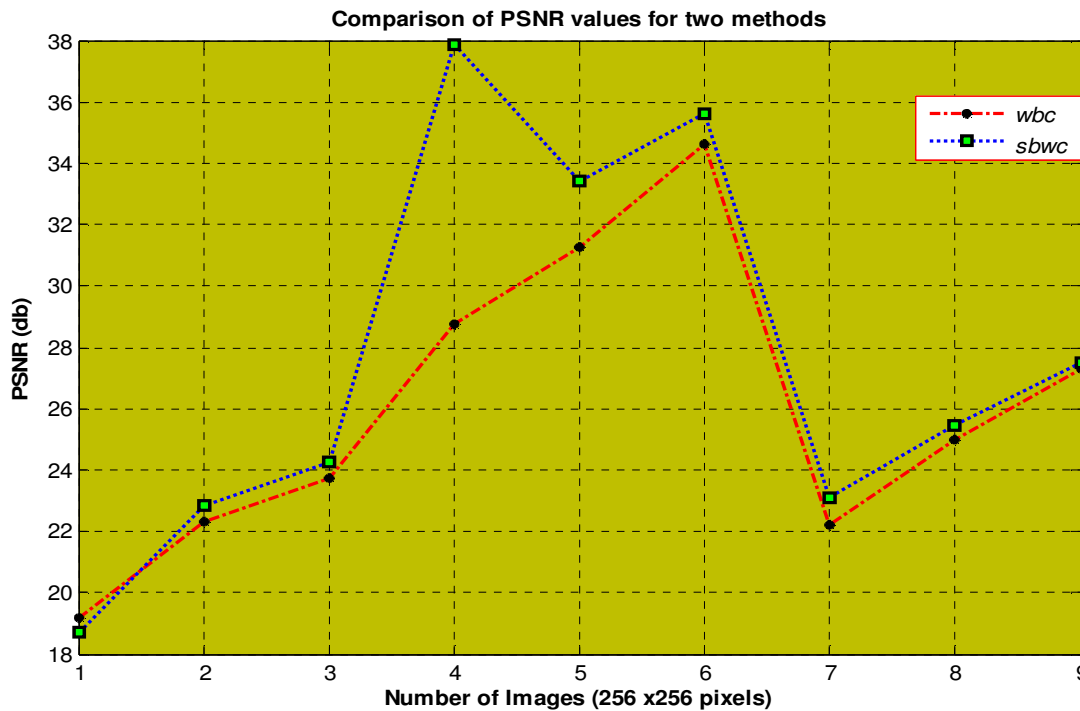


FIGURE 4: Comparison of PSNR for WBC and SBWC

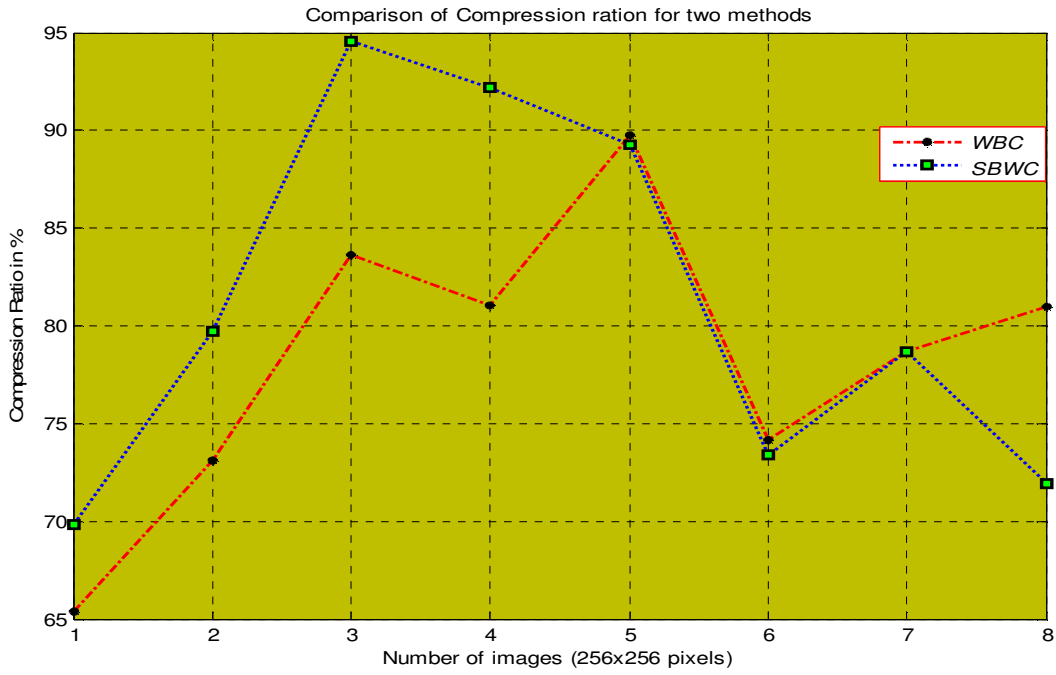


FIGURE 5: Comparison of CR for WBC and SBWC

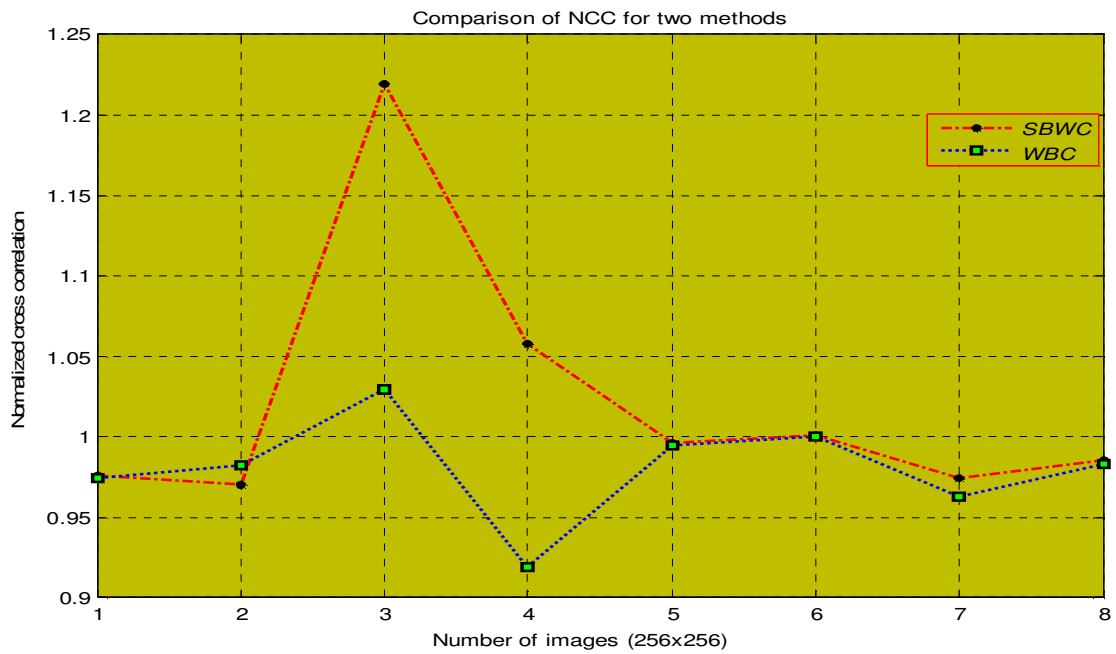


FIGURE 6: Comparison of NCC for WBC and SBWC

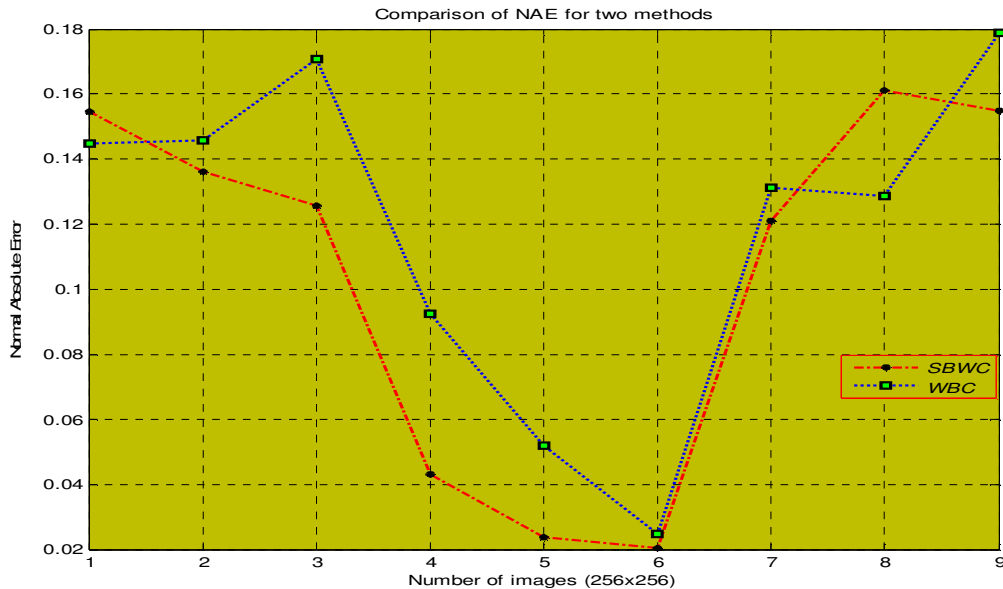


FIGURE 7: Comparison of NAE for WBC and SBWC

6. CONCLUSION

The objective specified is clearly evaluated by segmentation based image compression performance, in accounting compression ratio, PSNR and time needs. It is oriented toward the operational naval application intelligent imagery system. The images were evaluated quantitatively and qualitative assessments are done by metrics in order to establish the impact of wavelet and the need of segmentation based approach for image compression. The experimental analysis is mainly performed on SAR data from various satellite projections. Thus, images were processed in two different ways, analysis shows SBWC makes images smoother and conserves object edges without affecting the resolution. In ending, the Segmentation based wavelet compression technique allows output as expertly compressed image, with different operative importance without changing the pixel accuracy and image resolution or size. The analysis presented here will prove useful in studies of nonstationarity in time series, and the addition of statistical significance tests will improve the quantitative nature of wavelet analysis.

7. REFERENCES

- [1] Birgir Bjorn Saevarsson, Johannes R. Sveinsson and Jon Atli Benediktsson "Combined Wavelet and Curvelet Denoising of SAR Images" *Proceedings of IEEE 2004*.
- [2] A. Bruce and H. Gao. .Applied Wavelet Analysis with S-Plus.. Springer .Verlag New York, Inc. 1996.
- [3] S. G. Chang, B Yu and M Vetterli. .Adaptive Wavelet Thresholding for image Denoising and Compression.. IEEE Transactions on Image Processing, Vol. 9, No. 9, September 2000.
- [4]. Gersho A., Gray M. " Vector quantization and Signal compression", Kluwer Academic Publishers, Boston, 1992.
- [5] M. Grgic, M. Ravnjak, and B. Zovko-Cihlar, "Filter comparison in wavelet transform of still mages," in *Proc. IEEE Int. Symp. Industrial Electronics, ISIE'99*, Bled, Slovenia, pp. 105–110. 1999,

- [6] Guozhong Chen,Xingzhao Liu “Wavelet-Based Despeckling SAR Images Using Neighbouring Wavelet Coefficients.” *Proceedings of IEEE 2005*.
- [7] Guozhong Chen,Xingzhao Liu “An Improved Wavelet-based Method for SAR Images Denoising Using Data Fusion Technique”. *Proceedings of IEEE 2006*.
- [8]] J. Lu, V. R. Algazi, and R. R. Estes, “Comparative study of wavelet image coders,” *Opt. Eng.*, vol. 35, pp. 2605–2619, Sept. 1996.
- [9] Mario Mastriani “New Wavelet-based Superresolution Algorithm for Speckle Reduction in SAR Images” *IJCS volume 1 number 4, 2006*.
- [10] Mulcahy, Colm. .Image compression using the Haar wavelet transform.. Spelman Science and Math Journal. Found at: <http://www.spelman.edu/~colm/wav.html>.
- [11] Sonka, M. Hiaual, V. Boyle, R. Image Processing, Analysis and Machine Vision, 2nd edition. Brooks/Cole Publishing Company.
- [12] William B., Joan L, “Still Image Data Compression Standard”, Van Nostrand Reinhold, New York,1992
- [13] www.mdpi.com/journal/sensors *Article* Haiyan Li 1,2 , Yijun He 1,* and Wenguang Wang Improving Ship Detection with Polarimetric SAR based on Convolution between Co-polarization Channels.
- [14] Zhaohui Zeng and Ian Cumming,” SAR Image Compression Based on the Discrete Wavelet Transform”, Presented at the Fourth International Conference on Signal Processing ICSP'98, Beijing, China, October 12-16, 1998.
- [15] T. Chang and C. Kuo, "Texture Analysis and Classification with Tree-Structured Wavelet Transform", *IEEE Trans. Image Processing*, Vol. 2, No. 4, pp. 429-441, October 1993.
- [16] Chang.S and B.Yu, “Spatially adaptive wavelet thresholding with context modeling for image denoising.*IEEE trans Image processing* 9(9):533-539 ,2000.
- [17] S. Arivazhagan and L. Ganesan. Texture Segmentation Using Wavelet Transform. *Pattern Recognition Letters*, 24(16):3197–3203, December 2003.
- [18] M. G. Mostafa, T. F. Gharib, and coll. Medical Image Segmentation Using a Wavelet-Based Multiresolution EM Algorithm. *IEEE International Conference on Industrial Electronics Technology & Automation*, December 2001.

A Parametric Approach to Gait Signature Extraction for Human Motion Identification

Mohamed Rafi

*Dept. of Computer Science and Engineering
HMS Institute of Technology
Tumkur, Karnataka, India*

mdrafi2km@yahoo.com

Md. Ekramul Hamid

*Department of Computer Network Engineering
College of Computer Science
King Khalid University
Abha, Kingdom of Saudi Arabia*

ekram_hamid@yahoo.com

Mohamed Samiulla Khan

*Department of Computer Engineering
College of Computer Science
King Khalid University
Abha, Kingdom of Saudi Arabia*

mdsamiulla@gmail.com

R.S.D Wahidabanu

*Department. of E&C
Government college of Engg. Salem,
Tamil Nadu, India.*

drwahidabanu@gmail.com

Abstract

The extraction and analysis of human gait characteristics using image sequences are currently an intense area of research. Identifying individuals using biometric methods has recently gained growing interest from computer vision researchers for security purposes at places like airport, banks etc. Gait recognition aims essentially to address this problem by identifying people at a distance based on the way they walk i.e., by tracking a number of feature points or gait signatures. We describe a new model-based feature extraction analysis is presented using Hough transform technique that helps to read the essential parameters used to generate gait signatures that automatically extracts and describes human gait for recognition. In the preprocessing steps, the picture frames taken from video sequences are given as input to Canny edge detection algorithm which helps to detect edges of the image by extracting foreground from background also it reduces the noise using Gaussian filter. The output from edge detection is given as input to the Hough transform. Using the Hough transform image, a clear line based model is designed to extract gait signatures. A major difficulty of the existing gait signature extraction methods are the good tracking the requisite feature points. In the proposed work, we have used five parameters to successfully extract the gait signatures. It is observed that when the camera is placed at 90 and 270 degrees, all the parameters used in the proposed work are clearly visible. The efficiency of the model is tested on a variety of body position and stride parameters recovered in different viewing conditions on a database consisting of 20 subjects walking at both an angled and frontal-parallel view with respect to the camera, both indoors and outdoors and find the method to be highly successful. The test results show good clarity rates, with a high level of confidence and it is suggested that the algorithm reported here could form the basis of a robust system for monitoring of gait.

Keywords: Parametric Approach, Biometric, Gait Signature Extraction, Hough Transform, Canny Edge Detection, Gaussian Filter

1. INTRODUCTION

Gait analysis is a challenging research topic and recently that identification from gait has received attention and has become an active area of computer vision. For biometrics research, gait is usually referred to include both body shape and dynamics, i.e. any information that can be extracted from the video of a walking person to robustly identify the person under various condition variations. The demand for automatic human identification system is strongly increasing and growing in many important applications, especially at a distance and it has recently gained great interest from the pattern recognition and computer vision researchers for it is widely used in many security-sensitive environments such as banks, parks and airports [1].

The extraction and analysis of human walking movements, or gait, has been an on going area of research since the advent of the still camera in 1896 [2]. There two areas dominate the field of gait research at the present. Clinical gait analysis focuses on collection of gait data in controlled environments using motion capture systems and Biometric goals of human gait analysis analyze an individual's gait in a variety of different areas and scenarios. Because of these limitations, gait analysis for use in biometric systems are largely based on visual data capture and analysis systems which process video of walking subjects in order to analyze gait. The limitations that are inherent in these techniques necessitate the use of sophisticated computer vision systems to generate parameters which describe the gait motion.

Although considerable research has been done in the above areas, very limited successful research has been done in the area of gait extraction. The first scientific article on animal walking gaits has been written 350BC by Aristotle [3]. He observed and described different walking gaits of bipeds and quadrupeds and analyzed why all animals have an even number of legs. Recognition approaches to gait were first developed in the early 1990s and were evaluated on smaller databases and showed promise. DARPA's Human ID at a Distance program [4] then collected a rich variety of data and developed a wide variety of technique and showed not only that gait could be extended to large databases and could handle covariate factors. Since the DARPA program, research has continued to extend and develop technique, with especial consideration of practical factors such as feature potency. In Silhouette Analysis-Based Gait Recognition for Human Identification [5] a combination of background subtraction procedure and a simple correspondence method is used to segment and track spatial silhouettes of an image, but this method generates more noise which leads to poor gait signature extraction. Therefore the rate of recognition is low. In gait recognition by symmetry analysis[6], the Generalized Symmetry Operator is used which locates features according to their symmetrical properties rather than relying on the borders of a shape or on general appearance and hence does not require the shape of an object to be known in advance. The evaluation was done by masking with a rectangular bar of different widths in each image frame of the test subject and at the same position. A recognition rates of 100% were obtained for bar size of 5 pixels. This suggests that recognition is likely to be adversely affected when a subject walks behind a vertically placed object. There are also certain other limitations, Mark Ruane Dawson[6], like the legs were not being tracked to a high enough standard for gait recognition. The segmentation process leads to a very crude model fitting procedure which in turn adversely affects the rate of recognition. In other method of gait recognition, the subjects in the video are always walking perpendicular to the camera[7]. This would not be the case in real life as people would be walking at all angles to the video camera. Using of fewer parameters for gait signatures is a major drawback which has to be addressed. Cunado et.al. [8] use the Hough transform to extract lines that represent legs. Then they perform Fourier analysis on the resulting joint-angles for gait recognition. Specifically, they weight the magnitude spectra by the phase, to distinguish between different subjects. A CCR of 90% was obtained with the leave one out strategy. However, the algorithm was tested on only 10 subjects. The majority of published approaches fall into the model fitting category. Some rely on first processing each frame independently and then using PCA [9] or HMM[10] to model the transitions from one frame to the next. In [11], Nash et al. proposed a new model-based technique to allow the automated determination of human gait characteristics. Their technique employs a parametric two-line model representing the lower limbs. To speed up the search of the parameter space, they used a genetic algorithm (GA) based implementation of the Velocity

Hough Transform (VHT) rather than an exhaustive search. Although their approach is promising, the accuracy of the estimated hip rotation patterns is still insufficient for biometric purposes. More recently, methods that rely on dense optical flow by J. Little [12] or self similarity plots computed via correlation of pairs of images as by C. BenAbdelkader [13] have been proposed. The main drawback of these appearance based approaches is that they are usually designed only for a specific viewpoint, usually front to-parallel. Furthermore guaranteeing robustness against clothing and illumination changes remains difficult even though much effort has been expended to this end, for example by processing silhouettes or binary masks instead of the image itself [13]. Among model-based methods, is the one by Yam et al. [14] in which leg motion is extracted by temporal template matching with a model defined by forced coupled oscillators. Individual signatures are then derived by Fourier analysis. Another good recent example of model-based gait recognition can be found in [15]. The gait signature is extracted by using Fourier series to describe the motion of the upper leg and by applying temporal evidence gathering techniques to extract the moving model from a sequence of images. However these techniques are still 2-D, which means that a near fronto-parallel view is assumed. The approach we propose can be viewed as an extension of this philosophy to full 3-D modeling by replacing the Fourier analysis by the fitting of our PCA-based motion.

The motivation behind this research is to find whether increase in number of gait signature can improve the recognition rate? Improvement over model fitting can give us better results? What factors affect gait recognition and to what extent? And what are the critical vision components affecting gait recognition from video? The objective of this paper is to explore the possibility of extracting a gait biometric from a sequence of images of a walking subject without using markers. Sophisticated computer vision techniques are developed, aimed to extract a gait signature that can be used for person recognition.

In the proposed study, pre-processing is performed on the images in the walking sequence using the Canny edge detector is applied to produce edge images of the input data. Essentially, the human gait model describes a moving line whose inclination is constrained by a periodic signal and velocity governed by some initial conditions and characteristics. Then the model parameters are extracted and the line images are produced by Hough transform for lines. Using video feeds from conventional cameras and without the use of special hardware, implicates the development of a marker less body motion capture system. Research in this domain is generally based on the articulated-models approach. Haritaoglu et al. [10] presented an efficient system capable of tracking 2D body motion using a single camera. Amos Y. Johnson[11] used a single camera with the viewing plane perpendicular to the ground plane, 18 subjects walked in an open indoor-space at two view angles: a 45° path (angle view) toward the camera, and a frontal-parallel path (side-view) in relation to the viewing plane of the camera. The side-view data was captured at two different depths, 3.9 meters and 8.3 meters from camera. These three viewing conditions are used to evaluate our multi-view technique. In this research, we use images captured at different views as the image captured from the frontal or perpendicular view does not give required signatures. Segmentation is done on the captured image in order to extract foreground from back ground using Canny edge detection technique, as the purpose of edge detection in general is to significantly reduce the amount of data in an image, while preserving the structural properties to be used for further image processing. In order to obtain the gait model the output of segmentation is processed using Hough transform, which is a technique that can be used to isolate features of a particular shape within a Hough transformed image.

2. GAIT SIGNATURE EXTRACTION: THE PROPOSED MODEL

In medical science, the goal of most gait research has been to classify the components of gait and produced standard movement patterns for normal people that were compared to the gait patterns for pathological patients. No statistical or mathematical analysis was performed on the collected data [16]. We propose a gait signature extraction model having the following steps- picture frame capture, Segmentation, feature Extraction which leads to gait signature identification which are shown in figure.1.

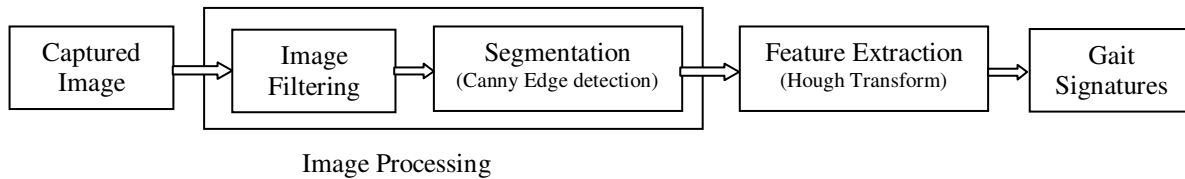


FIGURE 1: Components of the proposed model for Gait Signature Extraction System.

2.1 GAIT CAPTURING

At this step the subjects are asked to walk for capturing of gait. This is a very important step as the total result depends on the quality of the gait captured. So the care should be taken to see that the quality of gait capturing is maintained, this step includes video sequence and XML data store.

2.1.1. Video Sequence – the video sequence is the medium used for viewing the gait of an individual. Video will be taken in two side views(90 degree and 270 degree) so that we can get clear signature of each subject and the Video will be transformed into Frames and frames are stored in an intermediate format, bitmap or JPG, once being loaded into the program in order to allow process of the information within the program.

2.1.2. XML Data Store – the data store is used to record all of the captured data about each subject in XML format. Users gait data can either be stored in individual query files, or concatenated together to create a database file.

In our proposed research the following preprocessing steps are carried out before segmenting an Image

1. [Reading a RGB image]
2. [Converting an RGB image to Gray Scale]
3. [Converting Gray Scale Image to Indexed Image]

So, we can defined the captured image as two dimensional array of pixels $P = (P_x, P_y)$ and each pixel is an intensity I_p in 2D vector (B/W image). The indexed image is the input to the segmentation algorithm for further processing. The above preprocesses of an image is shown in figure. 2.

2.2. SEGMENTATION

In order to be able to perform analysis on the gait of the individuals caught on video the subject needs to be extracted from the video sequence. Image segmentation is used to separate dynamic objects such as people, which are part of the foreground, from the background of the image sequence. In computer vision, segmentation refers to the process of partitioning a digital image into multiple segments. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse. We use Image segmentation to locate an object and the boundaries of the object in images.

2.2.1 Edge Detection: Canny Edge Detector Algorithm

Segmentation by edge detection is a fundamental tool in image processing and computer vision, particularly in the areas of feature detection and feature extraction, which aim at identifying points in a digital image at which the image brightness changes sharply or more formally has discontinuities. The purpose of detecting sharp changes in image brightness is to capture important events and changes in properties of the world. In the ideal case, the result of applying an edge detector to an image may lead to a set of connected curves that indicate the boundaries of objects, the boundaries of surface markings as well as curves that correspond to discontinuities in surface orientation. Thus, applying an edge detection algorithm to an image may significantly reduce the amount of data to be processed and may therefore filter out information that may be regarded as less relevant, while preserving the important structural properties of an

image. The implementation of The Canny Edge Detection Algorithm runs in four separate steps [17]:

1. Image Smoothing:

Before doing any type of edge detection, it is important to filter the image to smooth out any noise picked up during image capture. This is essential because noise introduced into an edge detector can result in false edges output from the detector. The Canny edge detector uses a filter based on the first derivative of a Gaussian, the Gaussian Blur ξ has the form:

$$\xi[I]_p = \sum_{q \in S} G_\sigma(x, y) I_q \tag{1}$$

where σ is the size of the window, I_p and I_q are the values of intensity at pixel p and q , and the Gaussian function is defined as

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

Equation 1 indicates the filtered image at pixel p .

2. Finding gradients:

After smoothing the image and eliminating the noise, the next step is to find the edge strength by taking the gradient of the image. The edges should be marked where the gradients of the image has large magnitudes. The Canny algorithm basically finds edges where the grayscale intensity of the image changes the most. These areas are found by determining gradients of the image. First step is to approximate the gradient in the x- and y-direction respectively by applying the kernels. The gradient magnitudes or the edge strengths $\mathfrak{S}_s(x, y)$ can then be determined as an Euclidean distance measure by applying the law of Pythagoras given by equation

$$\mathfrak{S}_s(x, y) = \sqrt{\xi_x^2(x, y) + \xi_y^2(x, y)} \tag{2}$$

where, ξ_x and ξ_y are the estimated gradient in the form of 3x3 convolution masks in the x and y directions, respectively. The direction of the edges are determined as

$$\mathfrak{S}_o(x, y) = \tan^{-1}\left(\frac{\xi_y}{\xi_x}\right) \tag{3}$$

3. Non-maximum suppression:

Only local maxima should be marked as edges. The purpose of this step is to convert the “blurred” edges in the image of the gradient magnitudes to “sharp” edges. Basically this is done by preserving all local maxima in the gradient image, and deleting everything else. The non-maximum suppression is defined as:

$$\mathfrak{R}_N(x, y) = \begin{cases} \mathfrak{S}_s(x, y) & \text{if } \mathfrak{S}_s(x, y) > \mathfrak{S}_s(x', y') \& \mathfrak{S}_s(x, y) > \mathfrak{S}_s(x'', y'') \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $\mathfrak{S}_s(x', y')$ and $\mathfrak{S}_s(x'', y'')$ are the gradient magnitudes on both sides of edge at (x, y) in the direction of the gradient.



FIGURE 2: image of a walking person used in this study. [a] Original Image [b]. RGB to Grayscale [c] Grayscale to Indexed Image [d] Edge Detected Image.

4. Canny hysteresis thresholding:

Intensity gradients which are large are more likely to correspond to edges than if they are small. It is in most cases impossible to specify a threshold at which a given intensity gradient switches from corresponding to an edge into not doing so. Therefore Canny uses thresholding with hysteresis. For that we find out the local maximums that are true edges. We assume that true edges should have large strength and pixels belong to true edges are connected to each other. Now on the basis of the first assumption, we thresholded $\mathfrak{R}_N(x, y)$ using hysteresis algorithm. The algorithm uses two thresholds, τ_l and τ_h .

$$\mathfrak{R}_N(x, y) = \begin{cases} \mathfrak{R}_{Ns}(x, y) & \text{if } \mathfrak{R}_N(x, y) > \tau_h \\ \mathfrak{R}_{Nw}(x, y) & \text{if } \mathfrak{R}_N(x, y) \leq \tau_l \\ \mathfrak{R}_{Nc}(x, y) & \text{otherwise} \end{cases} \quad (5)$$

where $\mathfrak{R}_{Ns}(x, y)$, $\mathfrak{R}_{Nw}(x, y)$ and $\mathfrak{R}_{Nc}(x, y)$ are the strong, weak and candidate pixels at pixel (x, y) respectively. In each point of (x, y) , discard $\mathfrak{R}_{Nw}(x, y)$ and allow to pass $\mathfrak{R}_{Ns}(x, y)$. If the pixel is $\mathfrak{R}_{Nc}(x, y)$, then follow the chain of connected local maxima in both directions along the edge, as long as $\mathfrak{R}_N(x, y) > \tau_l$. Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge as shown in figure 2.

2.3. GAIT SIGNATURE EXTRACTION

Model based approaches to feature extraction, use a priori knowledge of the object, which is being searched for in the image scene. In this research, we are designing a line base model using Hough transform. When modeling the human body, there are various kinematical and physical constraints, we can place on the model which are realistic. The advantages of a model based approach are that evidence gathering techniques can be used across the whole image sequence before making a choice on the model fitting. Models can handle occlusion and noise better and offer the ability to derive gait signatures directly from model parameters i.e. variation in the inclination of the thigh. They also help to reduce the dimensionality needed to represent the data.

In this study, we model the human body using the Hough transform technique. Modeling the human body as a skeleton is a good representation due to the underlying skeleton of the physical body and the restrictions it creates. The model can be a 2- or 3-dimensional structural (or shape) model and motion model that lays the foundation for the extraction and tracking of a moving person. A gait signature that is unique to each person in the database is then derived from the extracted gait characteristics. One of the most important aspects in gait recognition is capturing accurately the positions of the legs, these are the best source for deriving a gait signature and will contain most of the variation in the subjects gait pattern. The legs of a human are usually modeled.

2.3.1. Hough Transform:

Hough (1962) proposed an interesting and computationally-efficient procedure for detecting lines in images- known as the Hough Transform. It is a technique which can be used to isolate features of a particular shape within an image. Because it requires that the desired features be specified in some parametric form, the classical Hough transform is most commonly used for the detection of regular curves such as lines, circles, ellipses, etc. A convenient equation for describing a set of lines uses *parametric* or *normal* notion in general form as:

$$x \cos \theta + y \sin \theta = r \quad (6)$$

where r is the length of a normal from the origin to this line and θ is the orientation of r with respect to the x-axis. The output from the Canny edge detection algorithm is given as input to the Hough transform to extract gait signatures in the form of lines. So for any point $\mathfrak{R}_{Ns}(x, y)$ on this line, r and θ are constant for $0 < \theta < 180$ is defined as

$$A(\theta, r) = +\{\mathfrak{R}_{Ns}(x, y) \in X \times Y | \mathfrak{R}_{Ns}(x) \cos \theta + \mathfrak{R}_{Ns}(y) \sin \theta + r = 0\} \quad (7)$$

where, $A(\theta, r)$ is an accumulator array and $X \times Y$ are the image domain. Then we search for the local maxima of $A(\theta, r)$ and which can be used to reconstruct the lines in the image. The edge detected image and the image after Hough transform is shown in figure 3.

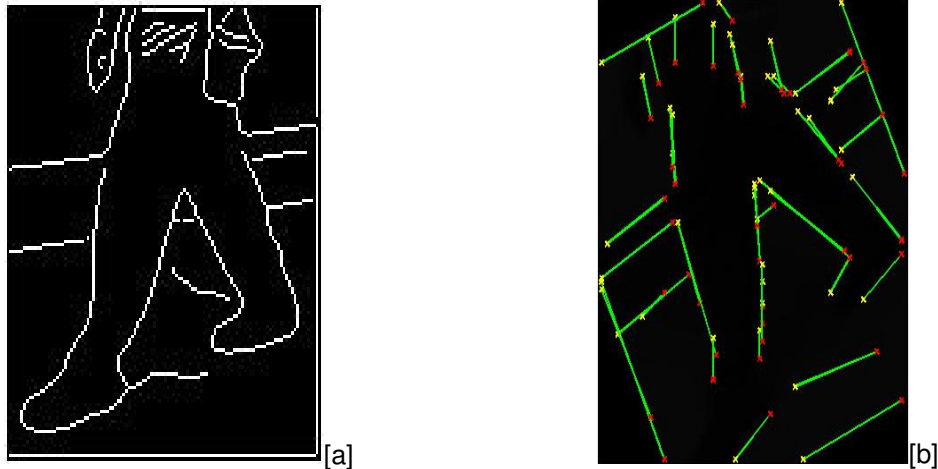


FIGURE 3: The Hough transform of an image. [a] Edge Detected Image and [b] Image after Hough Transform.

3. EXPERIMENTAL RESULTS AND DISCUSSION

3.1 Results

A video camera is used to collect data, and its output is recorded on a video recorder. The camera is situated with a plane normal to the subjects path in an environment with controlled illumination. Data collection is performed indoors, with lighting at a constant level. Subjects walked in front of a plain, static background. Each subject wore a special set of normal pant. In this way the camera-side leg could be distinguished visually from the other leg at all times. Fig. 3 shows an example image of a walking person used in this study. Each subject is asked walk past the camera a total of five times. From these five sequences, the first and last three are discarded and only the middle four sequences are used for experimentation. In the first few sequences the subject would be getting comfortable with the experiment, and in the last few the subject would be anxious to finish the experiment. As such, the middle four sequences were considered to offer the most consistent walking cycles. In all, 20 walking trials were completed, yielding the events for subsequent data analysis. This is a very important step as the total result depends on the quality of the gait captured. So the care should be taken to see that the quality of gait capturing is maintained, this step includes video sequence and XML data store. All post-processing and analysis was carried out off-line using the MATLAB programming environment.

We used an analytic approach, describing the legs and the motion of walking as a model. The human leg is modelled as two joined in series. The upper part is the thigh and is suspended between the hip and the knee. The lower part is the lower leg suspended from the knee to the ankle. We find that if all gait movements are considered, gait is unique and 5-7 distinct gait parameters are enough to identify a person. Using image processing techniques, lines representing legs in a sequence of images are extracted using the Hough transform. The inclination of the line representing the leg in each frame is collated to create the hip rotation pattern for the subject. Figure 4 shows the Hough transform of the image in Figure 3. Figures 4b-4g shows the different parameters used in this study.

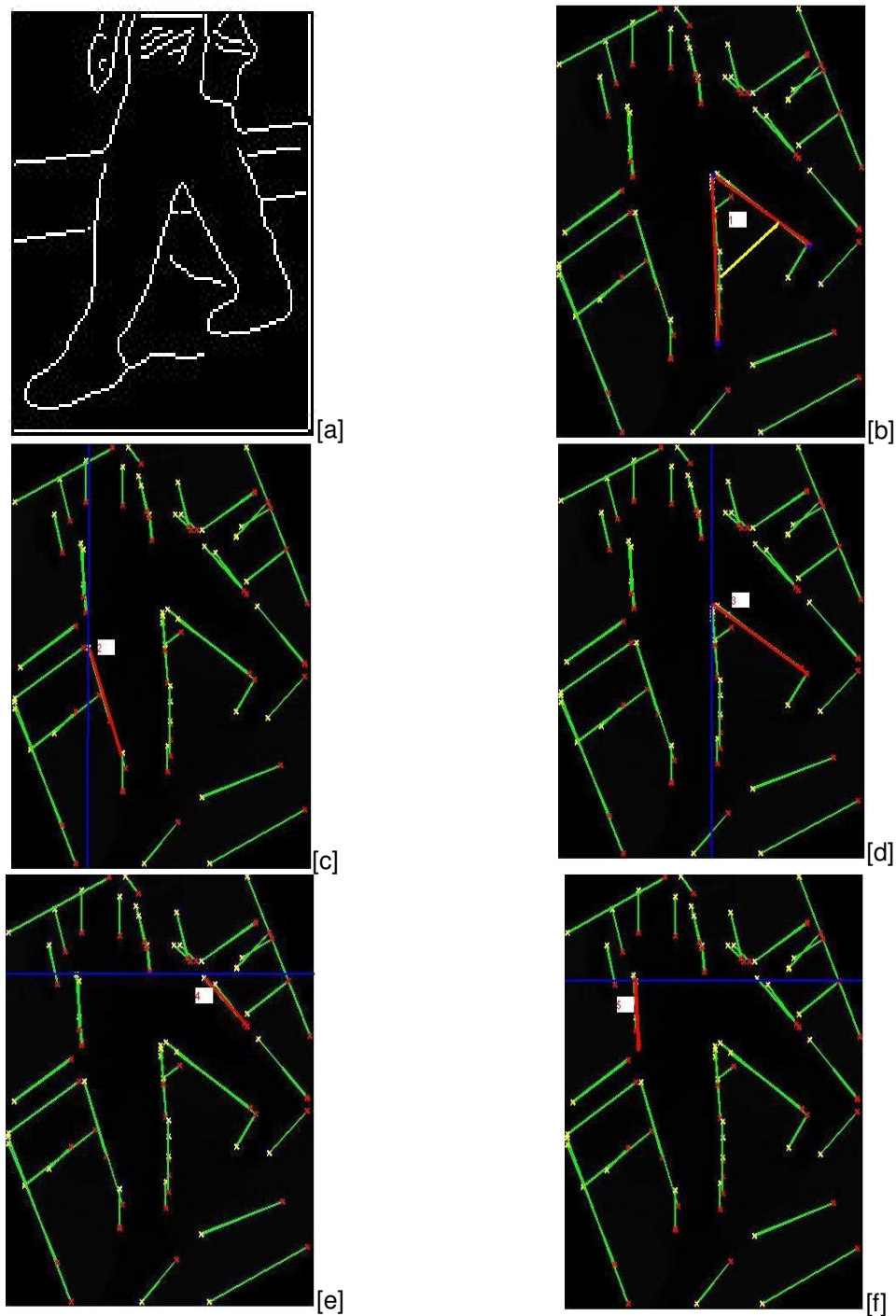


FIGURE 4: The Hough transformed image indicating different parameters. [a] edge detected image [b] distance between legs [c] right knee angle [d] left knee angle [e] left thigh angle [f] right thigh angle.

One of the most important aspects of this research is to extract the gait signatures for a successful recognition rate. Below table shows the number of parameters which are used to generate a gait signature for different view of a subject(90 degree and 270 degree).The attempts column shows how many persons were used to extract the signature. The success column shows how many of the subjects give successful gait signatures.

Parameter	Signature Parameter	Subjects	Clarity	Percentage %
1	Distance between legs	20	20	100
2	Right knee angle	20	18	90
3	Left knee angle	20	20	100
4	Left thigh angle	20	19	95
5	Right thigh angle	20	20	100

TABLE 1: Camera placed at 90 degree angle to the subject for frame 1.

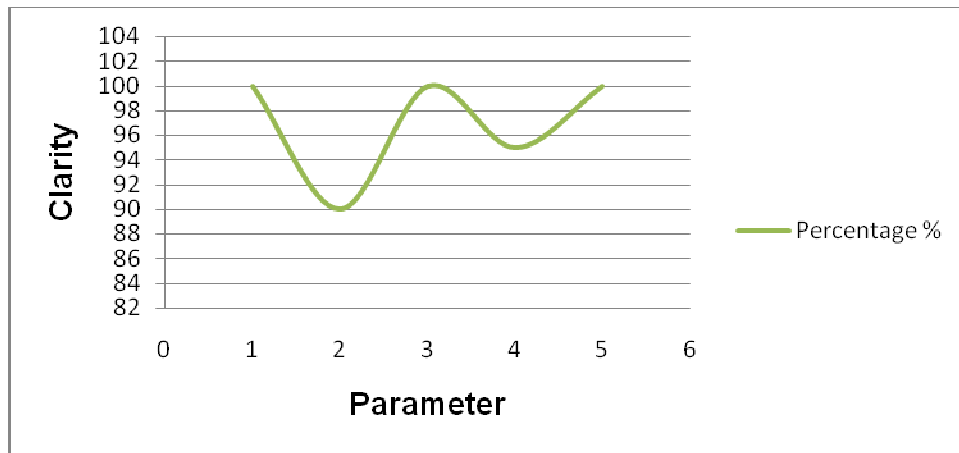


FIGURE 5: Graphical representation of clarity for frame 1, camera placed at 90 degree.

Parameter	Signature Parameter	Subjects	Clarity	Percentage %
1	Distance between legs	20	17	85
2	Right knee angle	20	18	90
3	Left knee angle	20	19	95
4	Left thigh angle	20	20	100
5	Right thigh angle	20	19	95

TABLE 2: Camera placed at 270 degree angle to the subject for frame 1.



FIGURE 6: Graphical representation of clarity for frame 1, camera placed at 270 degree.

Parameter.	Signature Parameter	Subjects	Clarity	Percentage %
1	Distance between legs	20	18	90
2	Right knee angle	20	18	90
3	Left knee angle	20	19	95
4	Left thigh angle	20	20	100
5	Right thigh angle	20	20	100

TABLE. 3: Camera placed at 90 degree angle to the subject for frame 2.

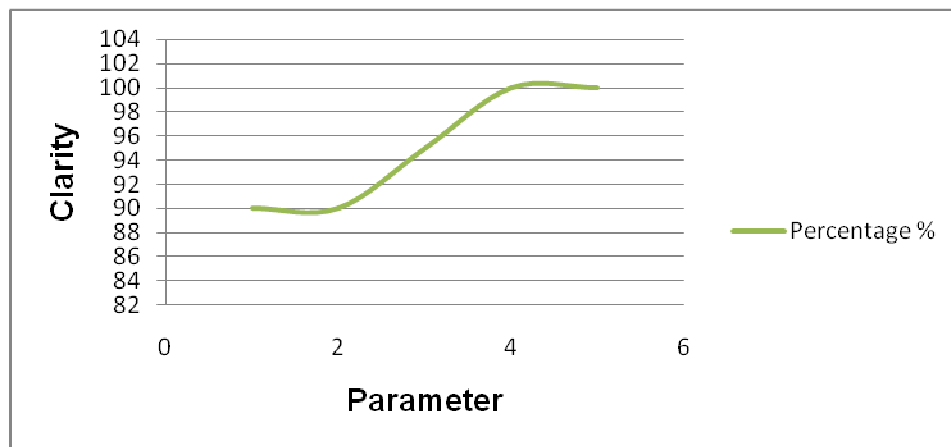


FIGURE 7: Graphical representation of clarity for frame 2, camera placed at 90 degree.

Parameter	Signature Parameter	Subjects	Clarity	Percentage %
1	Distance between legs	20	18	90
2	Right knee angle	20	18	90
3	Left knee angle	20	18	90
4	Left thigh angle	20	20	100
5	Right thigh angle	20	19	95

TABLE 4: Camera placed at 270 degree angle to the subject for frame 2.

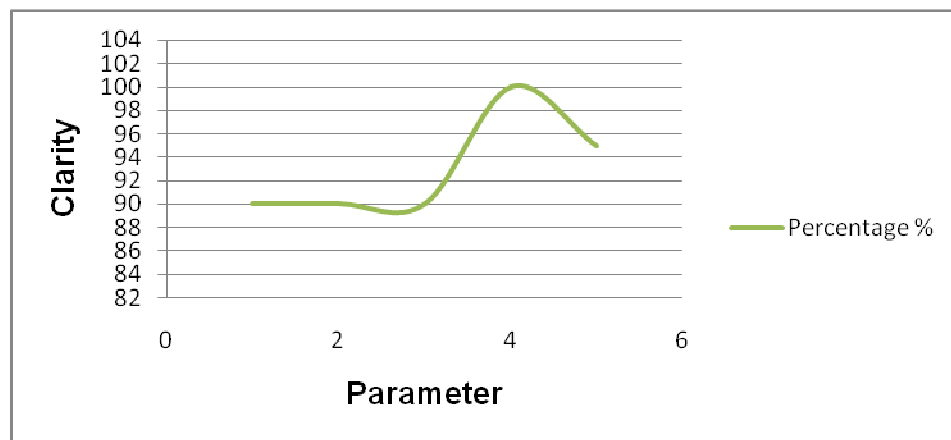


FIGURE 8. Graphical representation of clarity for frame 2, camera placed at 270 degree.

Parameter	Signature Parameter	Subjects	Clarity	Percentage %
1	Distance between legs	20	18	90
2	Right knee angle	20	19	95
3	Left knee angle	20	20	100
4	Left thigh angle	20	19	95
5	Right thigh angle	20	18	90

TABLE 5: Camera placed at 90 degree angle to the subject for frame 3.

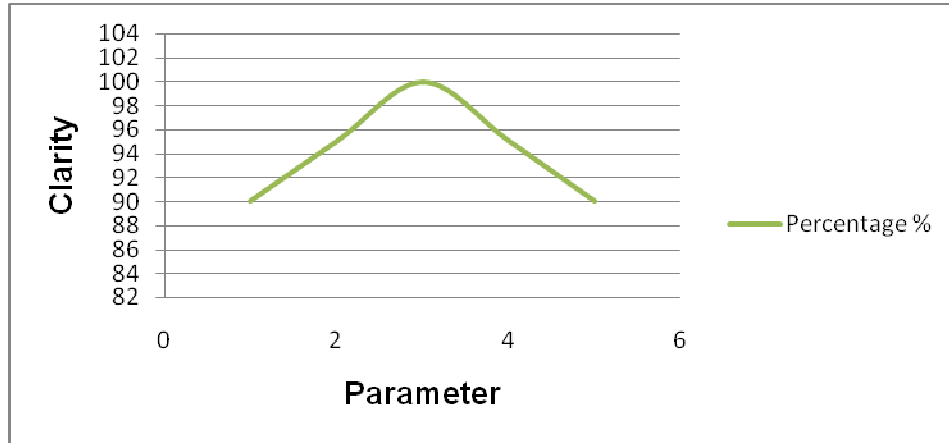


FIGURE 9: Graphical representation of clarity for frame 3, camera placed at 90 degree.

Parameter	Signature Parameter	Subjects	Clarity	Percentage %
1	Distance between legs	20	18	90
2	Right knee angle	20	17	85
3	Left knee angle	20	20	100
4	Left thigh angle	20	20	100
5	Right thigh angle	20	20	100

TABLE 6: Camera placed at 270 degree angle to the subject for frame 3.

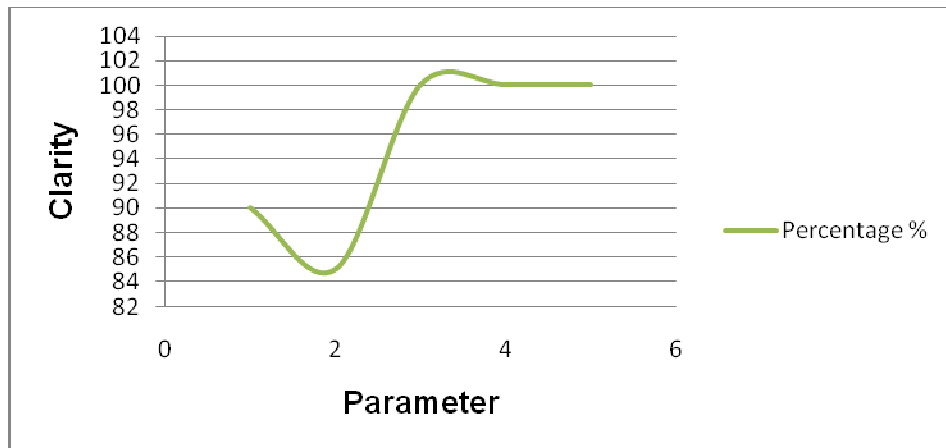


FIGURE 10. Graphical representation of clarity for frame 3, camera placed at 270 degree.

3.2 Discussions

While placing camera at 90 degrees and 270 degrees for frame 1, it is found that the clarity for the parameter distance between the legs is higher, the y axis is taken as the reference axis for the subject. Therefore this can be used to extract gait signatures for better recognition. It is also observed that the parameter right thigh angle can also be considered for extraction of gait signature.

While placing camera at 90 degrees and 270 degrees for frame 2, it is found that the clarity for the parameter right thigh angle is higher. Therefore this can be used to extract gait signatures for better recognition.

While placing camera at 90 degrees and 270 degrees for frame 3, it is found that the clarity for the parameter left thigh angle is higher. Therefore this can be used to extract gait signatures for better recognition. So the proposed study gives positive results in the use of gait as a biometric measure. A gait signature is extracted using computer vision techniques and produced a high correct classification rate on a small database of subjects.

4. CONCLUSION & FUTURE WORK

We were trying to prove a feature-based method could be used for gait recognition. It was shown that a gait signature based on parametric model gives an improved correct clarity rates for human gait recognition. This research has shown that gait signatures can be extracted in a better way by using Hough transform for extracting temporal features from image sequences. As a pre-processing step, Canny edge detector is applied to produce edge images of the input data. The experimental results for a database of 20 subjects showed that the proposed method for gait analysis could extract parameters for the human gait model with a high fidelity to the original image data. When the camera is placed at 90 and 270 degrees it is found that most parameters listed in the research are providing us clarity. Since the lines are clearly visible they can easily be labeled and the distance and angle between them can be measured accurately. The proposed research gives highest results if the camera is placed at 90 degrees to the subject and it is recommended that the subjects must be made to pass through a area which has a white background because it will help in getting a better gait signature extraction model. The research achieved 100 percent clarity if the parameters angle of left, right leg and distance between the legs are analyzed at 90 degree angle. Further experimentation will shows how these novel techniques can extract and describe gait, and results will be presented how they can be used to recognise people by their gait. Furthermore, the new technique extracts the gait signature automatically from the image sequence, without human intervention, one of the major aims of this work.

5. REFERENCES

- [1] Lu, A, Jiwen, Z. B. Erhu, "Gait recognition for human identification based on ICA and fuzzy SVM through multiple views fusion", School of Electrical and Electronic Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798, 25 July 2007.
- [2] J. Rose, and J. Gamble, Human Walking, 3rd edition, New York: Lippencott Williams and Wilkins, 2006.
- [3] Aristotle (350BC), On the Gait of Animals, Translated by A. S. L. Farquharson 2007.
- [4] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. Bowyer, "The humanID gait challenge problem: Data sets, performance and analysis", IEEE Trans. Pattern Anal. Mach. Intell., vol. 27, no. 2, pp. 162–177, Feb. 2005.
- [5] L. Wang, T. Tan, H. Ning, and W. Hu, "Silhouette Analysis-Based Gait Recognition for Human Identification", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 12, December 2003.

- [6] B. James, H. Acquah, M. S. Nixon, J. N. Carter, "Automatic gait recognition by symmetry analysis", Image, Speech and Intelligent Systems Group, Department of Electronics and Computer Science, University of Southampton, Southampton, S017 1BJ, United Kingdom.
- [7] X. Han, "Gait Recognition Considering Walking Direction", University of Rochester, USA, August 20, 2010.
- [8] D. Cunado, M. S. Nixon, and J. N. Carter, "Using gait as a biometric, via phase weighted magnitude spectra" first international conference on audio and video based biometric person authentication, pp. 95-102, 1997.
- [9] C. Fahn, M. Kuo and M. Hsieh, "A Human Gait Classification Method Based on Adaboost Techniques Using Velocity Moments and Silhouette Shapes", Next-Generation Applied Intelligence, Volume 5579/2009, 535-544, DOI: 10.1007/978-3-642-02568, 2009.
- [10] H. Murase and R. Sakai. "Moving object recognition in eigenspace representation: gait analysis and lip reading", Pattern Recognition Letters, 17:155–162, 1996.
- [11] J. M. Nash, J. N. Carter, and M. S. Nixon. "Extraction of Moving Articulated-Objects by Evidence Gathering", In Lewis and Nixon [Lewis98], pp. 609–18. September 1998.
- [12] Q. He and C. Debrunner. "Individual recognition from periodic activity using Hidden Markov Models.", In IEEE Workshop on Human Motion, Austin, Texas, December 2000.
- [13] S. Niyogi and E. Adelson. "Analyzing and recognizing walking figures in XYT.", in Conference on Computer Vision and Pattern Recognition, pages 469–474, 1994.
- [14] J. Little and J. Boyd. "Recognizing People by Their Gait: The Shape of Motion.", Videre, 1(2):1–32, 1986.
- [15] C. BenAbdelkader, R. Cutler, and L. Davis. "Motion-Based Recognition of People in EigenGait Space", In Automated Face and Gesture Recognition, pages 267–272, May 2002.
- [16] D. Cunado, M. S. Nixon, and J. N. Carter, Automatic extraction and description of human gait models for recognition purposes, Academic press, 2002.
- [17] D. A. Forsyth and J. Ponce, Computer Vision: A Modern Approach, Prentice-Hall, India, 2003.

An Experimental Study into Objective Quality Assessment of Watermarked Images

Anurag Mishra

*Department of Electronics,
Deendayal Upadhyay College, University of Delhi,
New Delhi 110015, India*

anurag_cse2003@yahoo.in

Aruna Jain

*Department of Computer Science,
Bharti College, University of Delhi,
New Delhi 110058, India*

arunajain1969@yahoo.co.in

Manish Narwaria

*School of Computer Engineering,
Nanyang Technological University,
Singapore 639798, Singapore*

mani0018@ntu.edu.sg

Charu Agarwal

*Department of Computer Science,
University of Delhi,
New Delhi 110007, India*

agarwalcharu2@rediffmail.com

Abstract

In this paper, we study the quality assessment of watermarked and attacked images using extensive experiments and related analysis. The process of watermarking usually leads to loss of visual quality and therefore it is crucial to estimate the extent of quality degradation and its perceived impact. To this end, we have analyzed the performance of 4 image quality assessment (IQA) metrics – Structural Similarity Index (SSIM), Singular Value Decomposition Metric (M-SVD), Image Quality Score (IQS) and PSNR on watermarked and attacked images. The watermarked images are obtained by using three different schemes viz., (1) DCT based random number sequence watermarking, (2) DWT based random number sequence watermarking and (3) RBF Neural Network based watermarking. The signed images are attacked by using five different image processing operations. The similarity parameter $SIM(X, X^*)$ is also computed for signed and attacked images after extracting watermarks and its computed values are correlated with other results. We observe that the metrics behave identically in case of all the three watermarking schemes. An important conclusion of our study is that PSNR is not a suitable metric for IQA as it does not correlate well with the human visual system's (HVS) perception. It is also found that the M-SVD scatters significantly after embedding the watermark and after attacks as compared to SSIM and IQS. Therefore, it is a less effective quality assessment metric for watermarked and attacked images. In contrast to PSNR and M-SVD, SSIM and IQS exhibit more stable and consistent performance. Their comparison further reveals that except for the case of counterclockwise rotation, IQS relatively scatters less for all other four attacks used in this work. It is concluded that IQS is comparatively more suitable for quality assessment of signed and attacked images.

Keywords: Digital Image Watermarking, Image Quality Assessment, PSNR, M-SVD, SSIM, IQS.

1. INTRODUCTION

Digital watermarking of images can be primarily classified into two categories: Perceptible and Imperceptible. These can be further classified as fragile or robust, depending upon the application. A watermarking technique can be perceptible and fragile on one hand and imperceptible and robust on the other. The criteria of imperceptibility and robustness are sometimes mutually exclusive. The watermark should be such that it is robustly embedded within the cover image and at the same time it should not degrade the perceptible quality of the signed image. Both these requirements have to be fulfilled in a well optimized watermarking algorithm. Precisely due to this reason, the digital watermarking of images is presently perceived as an optimization problem that can be solved using soft computing tools [1], [6]-[8].

Another very important requirement of a good watermarking technique is that it should not affect the perceived visual quality of the processed image significantly. That is, the modifications caused in the image due to the watermarking operation should be such that they are imperceptible to the Human Visual System (HVS). To this end, various researchers have published different schemes in which they attempt to model the HVS in order to watermark the cover images [2], [9]-[10]. Such watermarking schemes primarily take into account three different characteristics of the HVS namely luminance, contrast sensitivity and edge sensitivity. It has been established now that if the watermarking algorithm is implemented using HVS model, the objective quality score of the signed images is found to be higher. At the same time, if the watermark embedding procedure is well optimized, the robustness criterion is also fulfilled. In such a situation, the watermark embedding algorithm is the best optimized one.

Quality assessment of watermarked images is important since watermarking in general is a popular field for multimedia security related applications. Most studies in the existing literature have not focused on quality assessment of watermarked images. Conventionally, two image quality assessment (IQA) metrics – the Mean Squared Error (MSE) and its relative Peak Signal to Noise Ratio (PSNR) have been widely used and reported in literature [2], [6]. The PSNR is calculated using the MSE. Both these metrics are simple in terms of mathematical computation – both complexity and time. However, the physical meaning of MSE is simply the quantified assessment of the error in terms of difference between the original and distorted images. Therefore, the MSE and PSNR are not found to correlate well with the perception of the HVS. It is due to this reason that more advanced quality assessment models should be developed to quantify the quality of images in general and watermarked and attacked images in particular. Aziz and Chouinard [11] have probably for the first time reported an experimental approach to delve into the issue of IQA for watermarked images. They have used three different metrics – PSNR, weighted PSNR (wPSNR) and Watson Perceptual Metric (WPM) based on Just Noticeable Difference (JND) parameter for watermarked images obtained by using DWT decomposition of the original images. They have also argued that PSNR does not model HVS well. The second metric wPSNR is a modified variant of PSNR and is able to identify the textured areas of the original and watermarked images. Only the third metric – WPM, based on Watson model [11] is found to correlate with HVS up to some extent. The Watson model relies on context masking and pooling to create a response similar to the human perception. However, it can be noted that WPM does not take structural properties of the original or distorted / watermarked images into account. Shnayderman et. al [12] have computed the M-SVD values for watermarked images obtained by implementing seven standard watermarking schemes. They have reported that their numerical measure computes the global error (M-SVD) with values that range from 0.385 to 16.152 depending on what is embedded within the cover image. Thus, it can be noted that M-SVD values tend to scatter from minimum to maximum by introducing a small perturbation within the original images. Both groups [11]-[12] have not extended their work beyond watermark embedding and quality assessment of signed images. Recently, a new quality assessment metric CPA2 has been proposed specifically for watermarked images [13]. However, in this case also, the authors do not go beyond developing a new metric and ranking of only watermark embedding algorithms on the basis of computed perceptibility of signed images. Thus, overall, quality assessment of watermarked images remains a relatively uninvestigated area and calls for further investigations.

2. QUALITY ASSESSMENT METRICS USED IN IMAGE PROCESSING

As mentioned earlier, MSE/PSNR is the standard and widely used metric for quality assessment. Physically, it represents the error due to difference between the original and distorted images. Mathematically it is given, for a grayscale image wherein the intensity range of pixels is 0-255, by Eqn. 1.

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \| I(i, j) - K(i, j) \|^2 \quad (1)$$

where $I(i, j)$ and $K(i, j)$ are intensity values at pixel locations (i, j) in the original and distorted images of size $m \times n$ respectively. Lower the numerical value of MSE, better is the image quality. As the image is considered as a signal represented in spatial domain, any amount of noise or distortion may be found to affect its quality significantly. After adding noise to the image, the assessment of the ratio of signal component to noise component is done using PSNR. This parameter makes use of MSE and gives an assessment up to what extent the peak noise is present within the signal (image). Mathematically, it is given by Eqn. 2.

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_i^2}{MSE} \right) = 20 \cdot \log_{10} \left(\frac{MAX_i}{\sqrt{MSE}} \right) \quad (2)$$

where MAX_i is the maximum intensity of the pixels of the image and MSE is the mean square error for the pixels.

To overcome the aforementioned limitations of MSE/PSNR, substantial research effort has been put in recent years to devise their alternative [3]-[4], [14]. Some IQA metrics purely use the structural properties of the images [3], while others do not depend on them [4]. They rather make use of the singular values of the image matrix and quality of the image is evaluated on the basis of the difference in singular values of the original and perturbed images [4]. There are other models which make use of both - the structural properties of the images as well as their singular values [14]. Wang and Bovik [3] proposed a well cited image quality assessment metric known as Universal Image Quality (UIQ) and also its improved variant known as Structural Similarity Index Metric (SSIM). They have contended that in case of the SSIM model, image distortion is perceived as a combination of three different factors - loss of correlation, loss of mean distortion and loss of variance distortion. The SSIM is based on the concept that the HVS extracts structural information from the viewing field and therefore, a measurement of structural distortion can yield a good approximation of perceived image distortion. Mathematically, SSIM is given by Eqn. 3.

$$SSIM = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (3)$$

where $\mu_x = \frac{1}{n} \sum_{i=1}^n x_i$ and $\mu_y = \frac{1}{n} \sum_{i=1}^n y_i$ are mean intensity or luminance component of image signals x and y respectively. $C_1=(K_1L)^2$ and $C_2=(K_2L)^2$ are constants with L being the dynamic range of the grayscale image (0-255) and $K_1 \ll 1$ and $K_2 \ll 1$ being small constants. Besides this,

$$\sigma_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_x)^2, \quad \sigma_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \mu_y)^2 \quad \text{and}$$

$$\sigma_{xy}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y) \quad (4)$$

In Eqn. 4, σ_x and σ_y are signal contrast given by standard deviation for x and y respectively and is used to estimate contrast comparison for SSIM. The structure of the two images is associated

with the correlation (inner product) σ_{xy} between two unit vectors $(x_i - \mu_x)/\sigma_x$ and $(y_i - \mu_y)/\sigma_y$ lying in the image hyper plane.

Shnayderman et. al. [4] have applied the Singular Value Decomposition (SVD) to images to evaluate the image quality and named their metric as M-SVD. In this method, the image quality is evaluated by quantifying the difference of singular values between the original and distorted images. However, in this case, the singular vectors are completely ignored. In other words, by using singular values for quality evaluation, only the luminance factor is taken into account and the structural information contained in the singular vectors is not used. The numerical measure M-SVD is derived from the graphical measure given by Eqn. 5. It computes the global error expressed as a single numerical value depending on the distortion type as given by Eqn. 6.

$$\text{M-SVD} = \frac{\sum_{j=1}^B |D_j - D_{mid}|}{B} \quad (5)$$

where for each j^{th} block we calculate

$$D_j = \sqrt{\sum_{i=1}^b \{\sigma_i - \sigma_i^{(p)}\}^2} \quad (6)$$

where b defines the block size (in an image with size of $r \times c$), $B = r/b \times c/b$, and D_{mid} represents the midpoint of the sorted D_j 's. For the experimental results reported in this paper, we have used the commonly used block size of 8×8 i.e. $b=8$.

Narwaria and Lin [14] have argued that the mathematics of SVD provides a more general and characteristic view on changes of an image and structural information is crucial in image quality prediction. They have quantified the structural distortions in images using singular vectors. In their paper, they have further argued that the singular vectors denote clearer physical meaning for representing structural information, in comparison with the existing measures [3]-[4]. According to them, the changes in the singular vectors are related to those in the singular values which primarily reflect the image luminance only. Therefore, their proposed metric is capable of predicting visual quality in more general situations (i.e., with different types of distortions) as it takes into account singular vectors after decomposing the image matrix using SVD. The image quality score (IQS) is computed using the Minkowski Metric (MM) summation of the errors in singular vectors. IQA is also found to be scalable due to the ordering of singular vectors in the order of their perceptual significance. Moreover, the IQS has been shown to be an effective IQA metric demonstrating good prediction performance on a large number of distorted images. As said earlier, the numerical measure for change in \mathbf{U} and \mathbf{V} is expressed using Minkowski metric as

$$q_s = \left\{ \sum_{j=1}^l \gamma_j^m \right\}^{1/m} \quad (7)$$

where $m (>1)$ is a control parameter. A larger value of m puts more emphasis on large γ_j values. Narwaria and Lin [10] used $m=2$ for their practical computations. The final image quality score or IQS is determined using logarithmic scale. This is done to tackle the high dynamic range of q_s . Thus, we have

$$\text{IQS} = \ln(1 + q_s) \quad (8)$$

where the inclusion of constant 1 avoids the infinite value when q_s approaches 0. Since the first few singular vectors convey the major structural information, so Eqn. 7 is modified to obtain Eqn. 9 which will be finally used to compute IQS using Eqn. 8.

$$q_s = \left\{ \sum_{j=1}^K \gamma_j^m \right\}^{1/m} \quad (9)$$

where $K < t$ and K can be chosen depending on the computational cost metric performance. This shows the scalability of the IQS metric since we can use lesser number of singular vectors for quality prediction with reasonable and graceful degradation in prediction performance.

In the following sections, we report the experimental comparison of the performance of PSNR, SSIM, M-SVD and IQS for quality assessment of watermarked and attacked images. It can be noted that the four aforementioned metrics are full reference metrics i.e. they use the reference and the processed image for quality score computation. The watermarking of grayscale image Lena of size 256 x 256 pixels is done in low frequency DCT and DWT coefficients by three different methods. It has been established that the most robust watermarking can be achieved in case the watermark is embedded in the low frequency AC / DCT / DWT coefficients in transform domain of original image [5]. Moreover, any study on digital image watermarking is incomplete without executing image processing attacks over signed images and extraction of watermarks from signed as well as attacked images. Therefore, in the present work, five different image processing operations are executed to assess the visual quality of attacked images. Watermark extraction is performed both for the signed as well as attacked images of all three schemes. The similarity correlation parameter $SIM(X, X^*)$ is computed after extraction and correlated with other results for these images using Eqn 12.

3. EXPERIMENTAL WORK

This experimental work is classified in the following categories:

- (1) Watermarking of standard grayscale Lena image of size 256 x 256 pixels by using three different algorithms – (1) DCT based random number sequence watermarking, (2) DWT random number sequence watermarking and (3) Radial Basis Function Neural Network watermarking. In this work, three different schemes are chosen with an objective to observe and compare the performance of all three metrics on different watermark embedding algorithms. As a result of the watermarking operation, a set of three signed Lena images is obtained.
- (2) Applying five image processing attacks on these signed images as prescribed by StirMark standard. These attacks are (a) Counter clockwise rotation with 90° , (b) Gaussian Blur of radius = 1 unit, (c) Adding 10% Gaussian Noise to watermarked images, (d) Median Filtering (Filtering Aperture = 3) and (d) JPEG compression with quality factor $Q=0.9$ or compression ratio = 10%.
- (3) The watermarked and attacked images are subject to quality assessment by PSNR and three other metrics – SSIM, M-SVD and Image Quality Score (IQS). The results of the three metrics obtained for signed and attacked images are compared for their relative scattering before and after execution of attacks. To the best of our knowledge, any such experiment, the analysis of the data obtained thereafter and its comparison on watermarked and attacked images on the basis of these IQA metrics is never reported in literature.
- (4) Finally, watermark extraction from signed and attacked images of all three embedding schemes is performed. Similarity correlation parameter $SIM(X, X^*)$ is further computed for the extracted random number sequences from these images using Eqn. 12. The values of IQA

metrics are compared and analyzed in the context of computed $SIM(X, X^*)$ values for these images.

3.1 DCT based Random Number Sequence Watermarking

The grayscale image Lena is first decomposed into 8×8 blocks in spatial domain. All the blocks are then transformed into frequency domain by applying DCT algorithm. All AC coefficients of these blocks are sorted in decreasing order of their contribution to the image. Thus, the AC coefficients are now arranged in the descending order – from highest to lowest, the highest contribute the most while the lowest contribute the least. We have selected first 1024 coefficients from the top of this sequence. These coefficients are ones which constitute the low frequency coefficients of the image. According to Cox et. al [4], robust watermarking can be implemented if the watermarks are embedded in the low frequency coefficients of the image. However, in such a case, the effect of watermark embedding may be significant in terms of perceptible quality of signed image. In the present case, the watermark is a sequence of 1024 random numbers normalized with $N(0, 1)$. These random numbers are combined with the selected 1024 AC coefficients by using Eqn. 10 [5].

$$v'_i = v_i(1.0 + \alpha x_i) \tag{10}$$

where v_i are the low frequency DCT coefficients, $\alpha = 0.1$ is the scaling parameter also known as watermarking strength and x_i is the normalized random number of the watermark sequence. As a result of the computation outlined in Eqn. 10, the modified DCT coefficients v'_i are obtained. These modified coefficients v'_i are then reinserted at their respective pixel locations within the image in the transform domain. This image is applied with Inverse DCT (IDCT) to re-transform it into spatial domain. The resulting image is the signed image. This process of embedding the normalized random number sequence as watermark within the grayscale cover image Lena is depicted in Fig. 1.

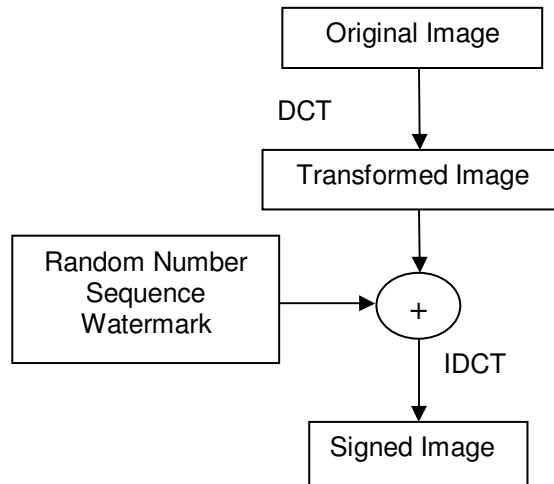


FIGURE 1: Block diagram of embedding random number sequence watermark within Lena using DCT

3.2 DWT based Random Number Sequence Watermarking

The second watermark embedding algorithm used in this work is based on computing three levels of DWT coefficients of grayscale Lena image. As the size of the cover image is 256×256 pixels, after three levels of wavelet decomposition using DWT using 'Haar filter', a block of size 32×32 coefficients is obtained as **LL3** region. These are 1024 low frequency coefficients obtained by applying DWT transform over the image. The same normalized random number sequence obtained previously is used as watermark to be embedded within these low frequency coefficients by using Eqn. 10. These modified coefficients will now be reinserted at their respective pixel

locations within the image in transform domain. The resulting image is then applied with Inverse DWT (IDWT) to obtain the signed image in the spatial domain. This process of embedding the random watermark sequence is depicted in Fig. 2.

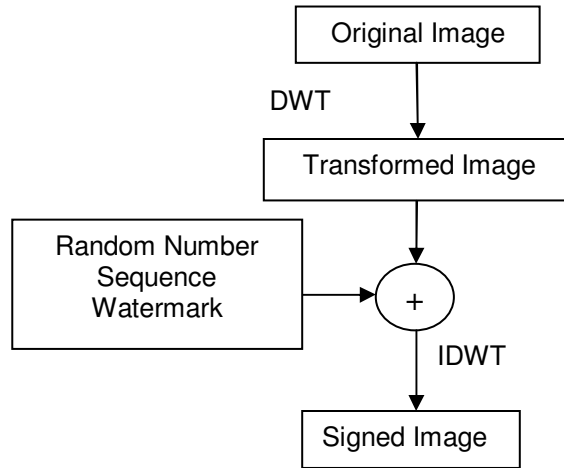


FIGURE 2: Block diagram of embedding random number sequence watermark within Lena using DWT

3.3 Radial Basis Function Neural Network (RBF – NN) based Watermarking

The third watermark embedding algorithm implemented in this work is close to one reported by Cheng-Ri Piao et. al [15]. They have proposed a new blind watermarking scheme in which a watermark is embedded within the image transformed by 4 - level DWT decomposition using ‘9/7 orthogonal filter’. Their method uses the HVS model and the Radial Basis Function Neural Network (RBF-NN). RBF is implemented while embedding and extracting the watermark by them. However, we have implemented only the watermark embedding component of their work. For this purpose, we have done a 3 – level DWT decomposition of the original image using ‘Haar filter’. The human visual system (HVS) model is used to determine the watermark insertion strength. The inserted watermark is the same normalized random number sequence used in previous two cases. The secret key used in the algorithm determines the beginning position of the image where the watermark is embedded. Listing 1 gives the embedding procedure used here which is similar to one proposed by Cheng-Ri Piao et. al [15].

Listing 1: RBF-NN Watermark Embedding Procedure

Step1: Transform an original image using the 3-level DWT transform. $C(i)$ is the LL3, LH3, HL3, and HH3 sub-band coefficients.

Step2: Select the beginning position of watermark embedding coefficient $C(i)$ using the secret key.

Step3: Quantize the DWT coefficient $C(i+key)$ by Q , as the input value of RBF then get the output $RBF(\text{round}(C(i+key)/Q))$

Step4: Embed the watermark according to the Eqn. 11 which uses the output value of the RBF neural network ($RBF(\text{round}(C(i+key)/Q))$) and Q .

$$c'_{i+key} = RBF \left[\text{Round} \left[\frac{c_{i+key}}{Q} \right] \right] + x_i \quad |x_i| \leq 7 \quad (11)$$

where x_i is the random sequence watermark, Q is a quantization value, and C'_i is the coefficient value when watermark is embedded. Then perform IDWT to get the watermarked image. Fig. 3 depicts the block diagram of this watermark embedding procedure.

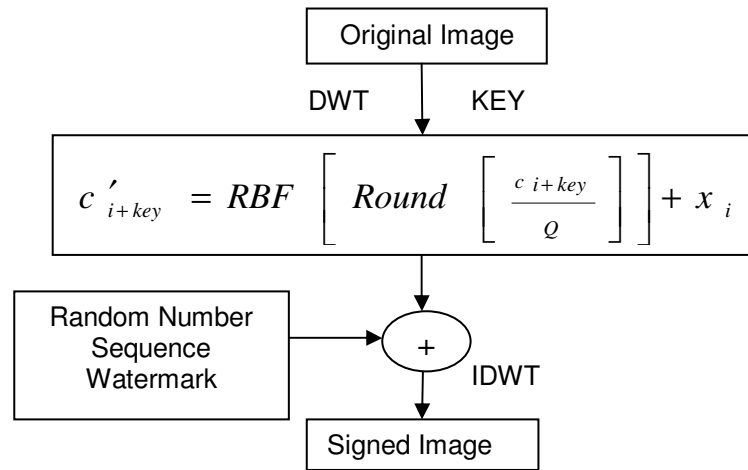


FIGURE 3: Watermark embedding procedure using RBF Neural Network

3.4 Executing Attacks Over Signed Images

As mentioned in Section III(2), the signed images obtained after embedding the watermarks in grayscale Lena image are subject to a five StirMark image processing attacks. These attacks are implemented by using MATLAB codes and Adobe Photoshop. These are aimed to verify the robustness and quality of signed images after attacks. Having implemented these attacks, we acquired a set of 3 signed images and 15 (5 attacks x 3 embedding procedures) attacked images.

3.5 Quality Assessment of Signed & Attacked Images

All these images are now checked for quality by using PSNR, SSIM, M-SVD and IQS. These results are compiled and tabulated in the Section IV.

3.6 Extraction of Watermarks and Computing SIM(X, X*) Correlation Parameter

The watermark is also extracted from all signed and attacked images of all three schemes. Similarity correlation parameter $SIM(X, X^*)$ is computed for the extracted watermarks using Eqn. 12. These computed values are analyzed and compared with other results of this work in Section IV and V.

$$SIM(X, X^*) = \frac{\sum_{i=1}^n (X - X^*)}{\sum_{i=1}^n \sqrt{(X^* \cdot X^*)}} \quad (12)$$

In Eqn. 12, X and X^* are respectively original and recovered watermark sequences.

4. RESULTS AND DISCUSSION

4.1 Quality Assessment of Watermarked Images

The original grayscale Lena image of size 256 x 256 is depicted in Fig. 4. Signed images obtained after embedding the random number sequence watermark using DCT based scheme, DWT based scheme and RBF-NN based scheme over this image are shown in Fig. 5-7 respectively.

Watermarking Scheme	PSNR	SSIM	M-SVD	Image Quality Score (IQS)
DCT Random Sequence Watermarking	43.1818	0.9983	5.5072	3.3449
DWT Random Sequence Watermarking	57.7806	0.9997	1.7861	3.4751
RBF-NN Watermarking	51.2441	0.9989	3.0530	3.4200

TABLE 1: PSNR, SSIM, M-SVD and Image Quality Score (IQS) for three embedding algorithms

It is clear that the perceptible quality of the signed images is very good. The quality is also objectively assessed and is tabulated in Table 1.

A careful observation of Table 1 indicates following points:

1. The quality of the signed image is the best after embedding the watermark in case of DWT based random number sequence algorithm while it is worst in case of DCT based random number sequence method, RBF-NN watermarking algorithm follows the DWT based method. This is indicated by all IQA metrics used in this work. Note that a higher M-SVD implies lower quality while higher values of IQS, SSIM and PSNR indicate higher quality. Thus, on the expected lines, we see that the M-SVD values decrease whereas PSNR, SSIM and IQS values increase as quality of the signed image goes from worst to the best.
2. Among the four quality assessment metrics used in this work, PSNR and M-SVD parameters tend to scatter significantly after introducing a small perturbation (watermark) within the cover image. As mentioned before, Shnayderman et. al [12] have reported that their numerical measure computes the global error with values that range from 0.385 to 16.152 depending on what is embedded within the cover image. We have, however, embedded the same random number sequence as watermark in all three watermarking schemes. In our case, the M-SVD values range from 1.7861 to 5.5072. This shows a significant scattering in M-SVD values for signed images which agrees with the results reported by Shnayderman et al [12]. On the other hand, the metrics that utilize structural information in images tend to be significantly stable. This can be explained since PSNR is nothing more than a simple average of pixel errors. Since the pixel error is computed by squaring (i.e. raising to power 2) the difference between the pixels, this may magnify/amplify small errors which are otherwise imperceptible to the human eye. A similar reasoning also explains the relatively poor performance of the M-SVD metric which also involves squared difference between the singular values of the original and watermarked image blocks. In contrast to this, metrics such as SSIM and IQS tend to judge the quality based on the overall structural changes and therefore tend to exhibit more stable and consistent performance. In view of the above, we believe that the SSIM and the IQS are more suitable candidates for the quality assessment of watermarked images.

We now further extend our experiment from simple watermark embedding within the cover image to implementing image processing attacks on the signed images obtained thereafter. The results of attacks are presented and analyzed in the next section.

4.2 Assessment of Watermarked Image Quality after Executing Image Processing Attacks

As mentioned in III(2), the attacks applied on the watermarked images are: (a) Counter clockwise rotation with 90^0 , (b) Gaussian Blur of radius = 1 unit, (c) Adding 10% Gaussian Noise to watermarked images, (d) Median Filtering (Filtering Aperture = 3) and (e) JPEG compression with quality factor $Q=0.9$ or compression ratio = 10%. Figs. 8-10(a-e) show five images each for three watermarking schemes after executing these attacks on images depicted in Fig. 5-7 respectively. We have computed IQA metrics for the three watermarking schemes after executing attacks in two different ways. In the first method, the test image is the attacked image while the reference image is the corresponding watermarked image. Table 2 shows the computed values of PSNR, SSIM, M-SVD and IQS obtained from test images of Figs. 8-10(a-e) and reference images of Figs. 5-7 respectively.

Watermarking Scheme	Image Processing Attack (Test Image)	PSNR (dB)	M-SVD	SSIM_INDEX	Image Quality Score (IQS)
DCT Watermarking Ref Image: Fig. 5	Rotation	11.4518	211.1825	0.2240	1.3949
	Gaussian Blur	29.1517	51.0299	0.8726	2.3687
	Gaussian Noise	30.6124	6.0670	0.7778	2.5441
	Median Filtering	30.9455	43.7245	0.8934	2.4200
	JPEG Compression	39.4671	2.6068	0.9675	2.7914
DWT Watermarking Ref Image: Fig. 6	Rotation	11.6649	207.6996	0.2284	1.4101
	Gaussian Blur	29.1995	50.8598	0.8724	2.3548
	Gaussian Noise	30.5060	6.1556	0.7726	2.5180
	Median Filtering	30.9802	43.6350	0.8931	2.3787
	JPEG Compression	39.4862	2.7126	0.9674	2.8063
RBF-NN Watermarking Ref Image: Fig. 7	Rotation	11.6541	209.5146	0.2288	1.4137
	Gaussian Blur	29.2000	50.9478	0.8724	2.3607
	Gaussian Noise	30.5586	6.4721	0.7743	2.5099
	Median Filtering	30.9803	43.6449	0.8933	2.3847
	JPEG Compression	39.4869	2.8492	0.9674	2.8035

TABLE 2: Quantified values of Quality Assessment parameters obtained from images of Fig. 8-10(a-e). Test Image: Attacked (Column II), Reference Images: Figs. 5-7 respectively for three schemes

On the other hand, the second method deals with the test images of Figs. 8-10(a-e) and reference image of Fig. 4. Table 3 shows the computed values of PSNR, SSIM, M-SVD and IQS obtained from test images of Figs. 8-10(a-e) and reference image of Fig. 4 for all three watermarking schemes.

Watermarking Scheme	Image Processing Attack (Test Image)	PSNR (dB)	M-SVD	SSIM_INDEX	Image Quality Score (IQS)
DCT Watermarking Ref Image: Fig. 4	Rotation	11.5536	3242.0	0.2264	1.4957
	Gaussian Blur	29.13	432.795	0.8727	2.3410
	Gaussian Noise	30.3837	198.7199	0.7757	2.5228
	Median Filtering	30.6867	285.6602	0.8931	2.3840
	Jpeg Compression	37.9451	71.3018	0.9661	2.7737
DWT Watermarking Ref Image: Fig. 4	Rotation	11.6546	3204.5	0.2289	1.5004
	Gaussian Blur	29.195	442.834	0.8723	2.3666
	Gaussian Noise	30.5492	193.5546	0.7741	2.5172
	Median Filtering	30.9722	294.7823	0.8931	2.3956
	Jpeg Compression	39.4215	59.0194	0.9671	2.7894
RBF-NN Watermarking Ref Image: Fig. 4	Rotation	11.6584	3206.4	0.2286	1.5043
	Gaussian Blur	29.1818	442.5634	0.8720	2.3606
	Gaussian Noise	30.4724	196.8129	0.7718	2.5222
	Median Filtering	30.9534	294.3198	0.8926	2.4037
	Jpeg Compression	39.2051	65.9834	0.9663	2.7833

TABLE 3: Quantified values of Quality Assessment parameters obtained from images of Fig. 8-10(a-e). Test Image: Attacked (Column II), Reference Image: Fig. 4 for all three schemes

4.3 Behavior of $SIM(X, X^*)$ Parameter for Signed and Attacked Images

After extraction, the recovered watermarks are compared with the original watermarks using similarity correlation parameter $SIM(X, X^*)$ given by Eqn. 12. Fig. 11(a-c) depict the $SIM(X, X^*)$ plots for the random number watermarks extracted from images of Fig. 5-7 respectively. The $SIM(X, X^*)$ values in these three cases are **18.9206**, **19.2357** and **19.0732** respectively and are tabulated in Table 4. These values give rise to two prominent observations – (1) that the extraction process from the signed images is quite successful and (2) that the numerical values indicate that the extraction out of DWT watermarking scheme is the best, followed by RBF-NN and DCT watermarking schemes. This is expected as we have already concluded in section 4.1 that DWT watermarking scheme is the best whereas the RBF-NN and DCT watermarking schemes come second and third respectively. The extraction of watermarks is also done from attacked images of Figs. 8-10(a-e) and their $SIM(X, X^*)$ values are also computed. Table 4 compiles the $SIM(X, X^*)$ values for the five attacks for all three watermarking schemes. The rotation attack is found to be the worst for all three schemes as there is hardly any watermark recovery from the attacked image (very small $SIM(X, X^*)$ value). However, the Jpeg compression gives the best results among the five chosen attacks for all three schemes. These $SIM(X, X^*)$

values are also found to be well correlated with the respective values of M-SVD, SSIM and IQS as given in Table 2 and 3. For the DCT watermarking scheme, in case of rotation attack as per Table 3, the M-SVD, SSIM and IQS values are 3242.0, 0.2264 and 1.4957 respectively while they are 71.3018, 0.9661 and 2.7737 respectively for Jpeg compression attack. We know that M-SVD values increase whereas SSIM and IQS values decrease with the degradation of image quality after executing attacks. Therefore, rotation brings in the most degraded effects (worst quality) within the signed images whereas Jpeg does not degrade the signed images up to that extent. This trend can be explicitly correlated with respective $SIM(X, X^*)$ values tabulated in Table 4. For ex., the $SIM(X, X^*)$ values are 1.4093 and 18.894 respectively for rotation and Jpeg compression

Watermarking Scheme	Image / Attack	$SIM(X, X^*)$ Correlation Parameter
DCT Watermarking	DCT Watermarked Image	18.9206
	Rotation	1.4093
	Gaussian Blur	9.1398
	Gaussian Noise	9.8807
	Median Filter	13.6321
	Jpeg Compression	18.8940
DWT Watermarking	DWT Watermarked Image	19.2357
	Rotation	0.6088
	Gaussian Blur	18.3738
	Gaussian Noise	17.3372
	Median Filter	18.5614
	Jpeg	19.2083
RBF-NN Watermarking	RBF-NN Watermarked Image	19.0732
	Rotation	1.1165
	Gaussian Blur	13.0912
	Gaussian Noise	13.0564
	Median Filter	15.3813
	Jpeg	19.0285

TABLE 6: Computed $SIM(X, X^*)$ values for signed and attacked images of all three schemes

attacks for the DCT watermarking scheme. A smaller $SIM(X, X^*)$ indicates that there is hardly any watermark within the cover image. Even if the presence of watermark inside the cover image is known (informed watermarking), it further indicates that the attack executed on the signed image has degraded the watermark up to the extent that it is impossible to recover it. Besides this, different attacks bring in different transformations within the signed image data, thereby resulting in variance of signed image quality after executing attacks. Therefore, after rotation, the quality of

the signed image is found to be degraded more as compared to Jpeg operation. This correlated behavior of computed values of the IQA metrics on one hand and $SIM(X, X^*)$ parameter on the other is established in this work in case of all three watermarking schemes.

4.4 Behavior of IQA Metrics after Executing Attacks on Signed Images

Results compiled in Table 2 clearly indicate that the PSNR and M-SVD values scatter significantly due to reasons outlined in Section IV (A). In fact, M-SVD is found to scatter even more significantly in comparison to PSNR after executing the image processing operations. This may be due to the fact that the attacks directly affect the singular values (luminance component) of the signed images. More the intensity of attack, greater is the scattering in M-SVD values. For ex., the counterclockwise rotation of 90° results in lowering of PSNR from 43.1818 dB to 11.4518 dB in DCT based watermarking scheme. In other words, rotation causes significant loss of quality (highest intensity of attack) in the signed image. This is also reflected by computing M-SVD (5.5072 to 211.1825) after rotation for the same scheme. A similar result is obtained in other two watermarking schemes also. However, in case of DCT based scheme, the scattering is even more prominent. On the other hand, as indicated by the results, SSIM and IQS tend to be more stable and consistent. In case of DCT based scheme, after executing counter clockwise rotation of 90, the SSIM is found to vary from 0.9983 to 0.2240 whereas the IQS varies from 3.3449 to 1.3949. We observe an identical behavior of SSIM and IQS in case of other two watermarking schemes also. In addition to that, SSIM and IQS behave identically in case of all five executed attacks spanning all three watermarking schemes used in this work. Narwaria and Lin [14] have compared their Image Quality Score (IQS) with M-SVD and SSIM and maintained that it performs better both in terms of monotonicity and scatters less compared to others. They have established this by developing scatter plots for all three metrics. We have compared the performance of the three metrics on the basis of their relative scattering before and after executing image processing attacks. For ex., in case of DCT based watermarking scheme, for rotation attack, the relative scattering shown by M-SVD is $211.1825/5.5072 = 38.346$ and by SSIM is $0.2240/0.9983 = 0.2243$ whereas the same for IQS is $1.3949/3.3449 = 0.4170$. We have observed similar behavior of all three metrics in case of other two watermarking schemes also. Table 5 gives a comparison of the relative scattering shown by M-SVD, SSIM and IQS for reference images of Figs 5-7. It is clear from Table 5 that in case of attacked images, the relative scattering of M-SVD values is the highest. We have also computed the three IQA metrics and their relative scattering using Fig. 4 as reference image. The values of M-SVD, SSIM and IQS obtained in this case are tabulated in Table 6. It is very interesting to note that the relative scattering of SSIM and IQS is similar to one which is compiled in Table 5 while that of M-SVD is very prominent – several multiples of the values given in Table 5. Thus, besides PSNR, the M-SVD may also not be suitable for assessing the quality of the attacked images. The poor performance of M-SVD is not really surprising since it uses singular values which are relatively less crucial for IQA. We are now left with two other metrics namely IQS and SSIM both of which extract the structural information from original and attacked/distorted images. A comparison between these two can be done on the basis of their relative scattering given in Table 5 and 6. We have observed that only in case of counterclockwise rotation, the relative scattering of IQS is more than that of SSIM. A possible reason for this could be that IQS is sensitive to the relative orientation of the image structure. In other words, the rotation causes a large change in the singular vectors of the processed (i.e. attacked) image. This possibly results in large error on using the dot product between the singular vectors of the original and distorted images. However, it is also found that SSIM is also quite sensitive to such rotational changes and it remains an interesting avenue for possible future work to design IQA metrics which are capable to handle such distortion. Nonetheless, for other four attacks, the IQS shows comparatively less relative scattering than SSIM. This behavior is consistently observed in case of all three watermarking schemes used in this work. As far as

Watermarking Scheme	Attack (Test Image)	Relative Scattering shown by		
		M-SVD	SSIM	Image Quality Score (IQS)
DCT Watermarking Ref Image: Fig. 5	Rotation	38.346	0.2243	0.4170
	Gaussian Blur	9.266	0.874	0.7081
	Gaussian Noise	1.1016	0.7791	0.7605
	Median Filtering	7.9395	0.8949	0.7234
	Jpeg Compression	0.4733	0.9691	0.8345
DWT Watermarking Ref Image: Fig. 6	Rotation	117.3	0.2288	0.4068
	Gaussian Blur	28.524	0.8726	0.6793
	Gaussian Noise	3.6235	0.7745	0.7222
	Median Filtering	24.435	0.8935	0.6862
	Jpeg Compression	1.5952	0.9676	0.8067
RBF-NN Watermarking Ref Image: Fig. 7	Rotation	68.03	0.2286	0.4123
	Gaussian Blur	16.658	0.8733	0.6885
	Gaussian Noise	2.016	0.7734	0.7362
	Median Filtering	14.292	0.894	0.6955
	Jpeg Compression	0.888	0.9684	0.8205

TABLE 5: Relative scattering of M-SVD, SSIM and IQS values for three watermarking schemes. Test Images: Figs. 8-10(a-e) given in column II, Reference Images: Figs. 5-7 respectively for three schemes

SSIM and IQS are concerned, although, no IQA metric used in this work is found to be perfectly suitable to cater all attacks, yet, we find that IQS formulated by Narwaria and Lin [14], is comparatively a more stable and consistent metric for quality assessment of signed and attacked images.

Watermarking Scheme	Attack (Test Image)	Relative Scattering shown by		
		M-SVD	SSIM	Image Quality Score (IQS)
DCT Watermarking Ref Image: Fig. 4	Rotation	588.68	0.2268	0.4471
	Gaussian Blur	78.590	0.8741	0.6998
	Gaussian Noise	36.083	0.7770	0.7542
	Median Filtering	51.8703	0.8946	0.7127
	Jpeg Compression	12.947	0.9677	0.8292
DWT Watermarking Ref Image: Fig. 4	Rotation	1794.13	0.2289	0.4317
	Gaussian Blur	247.933	0.8726	0.6810
	Gaussian Noise	108.367	0.7743	0.7243
	Median Filtering	165.042	0.8933	0.6893
	Jpeg Compression	33.043	0.9674	0.8026
RBF-NN Watermarking Ref Image: Fig. 4	Rotation	1050.24	0.2288	0.4398
	Gaussian Blur	144.9601	0.8729	0.6902
	Gaussian Noise	64.4654	0.7726	0.7374
	Median Filtering	96.4034	0.8935	0.7028
	Jpeg Compression	21.6126	0.9673	0.8138

TABLE 6: Relative scattering of M-SVD, SSIM and IQS values for three watermarking schemes. Test Images: Figs. 8-10(a-e) given in column II, Reference Images: Fig. 4 for all three watermarking schemes

5. CONCLUSIONS

In this paper, we have used three different watermarking schemes, viz., DCT based watermarking, DWT based watermarking and RBF-Neural Network based blind watermarking for assessment of quality of images obtained after embedding the watermark which is a random

number sequence of size 1024 using Cox's formula [5]. In addition to this, the signed images are subject to five different StirMark prescribed image processing attacks – Counterclockwise rotation of 90^0 , Gaussian Blur of radius = 1 unit, (c) Adding 10% Gaussian Noise to watermarked images, (d) Median Filtering (Filtering Aperture = 3) and (d) JPEG compression with quality factor $Q=0.9$ or compression ratio = 10%. The quality of the signed and attacked images is assessed by three widely reported IQA metrics besides PSNR. These are Singular Value Decomposition Metric (M-SVD), Structural Similarity Index (SSIM) and Image Quality Score (IQS). As, the PSNR does not correlate well with Human Visual System (HVS model) of perception, it is not found suitable for quality assessment of signed and attacked images. This is due to the fact that PSNR represents the squared difference of corresponding pixel values between the original and distorted images and is calculated using Mean Square Error (MSE) between them. The extraction of the watermarks from signed and attacked images is also done in this work. The similarity correlation parameter $SIM(X, X^*)$ is used to compare the embedded and recovered watermarks from these images. The computed results of the IQA metrics are found to be well correlated with those of $SIM(X, X^*)$ parameter in case of all three watermarking schemes and attacks. Further, a comparison between three other metrics besides PSNR is done by estimating the relative scattering of their values after executing attacks. Among them, M-SVD values scatter much more significantly after watermark embedding and after executing image processing attacks for all three watermarking schemes. This is due to the reason that M-SVD is based on the squared difference between the block wise singular values or luminance of the original and distorted images. In this case, the structural information contained in the singular vectors obtained by the SVD decomposition of image matrix is completely ignored. Due to this reason, M-SVD is not found to be a stable, consistent and reliable candidate for quality assessment of signed and attacked images. The relative scattering observed in case of other two metrics – SSIM and IQS is comparatively much less. Both these metrics extract structural information from the original / reference and distorted / attacked images. Between SSIM and IQS, it is found that except counterclockwise rotation of 90^0 , for all other four attacks used in this work, the Image Quality Score (IQS) is more stable and consistent than its other counterpart. It is, therefore, concluded that among the three IQA metrics used in this work besides PSNR, the Image Quality Score (IQS) given by Narwaria and Lin [14] is comparatively more suitable for quality assessment of the watermarked and attacked images.

6. REFERENCES

- [1] Hsiang-Cheh Huang, Yueh-Hong Chen and Ajith Abraham, "Optimized Watermarking Using Swarm-Based Bacterial Foraging", Journal of Information Hiding and Multimedia Signal processing, 1(1):51-58, 2010
- [2] Der-Chyuan Lou, Ming-Chiang Hu, and Jiang-Lung Liu, "Healthcare Image Watermarking Scheme Based on Human Visual Model and Back-Propagation Network", Journal of C.C.I.T, 37(1): 151-162, 2008
- [3] Zhou Wang, Alan C. Bovik and Hamid R. Sheikh, "Image Quality Assessment: From Error Measurement to Structural Similarity", IEEE Transactions on Image Processing, 13(1), January 2004
- [4] A Shnayderman, A Gusev and A M. Eskicioglu, "An SVD-Based Gray-Scale Image Quality Measure for Local and Global Assessment", IEEE Transactions on Image Processing, 15(2), February 2006
- [5] Ingemar J. Cox, Joe Kilian, F. Thomson Leighton, and Talal Shamooh, "Secure Spread Spectrum Watermarking for Multimedia", IEEE Transactions on Image Processing, 6(12):1673-1687,1997
- [6] Shieh, C., H. Huang, F. Wang and J. Pan, "Genetic Watermarking based on Transform Domain Techniques", Pattern Recognition Letters, vol. 37, pp. 555-565, 2004

- [7] Charu Agarwal and Anurag Mishra, “*A Novel Image Watermarking Technique using Fuzzy-BP Network*”, In the Proc. Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Darmstadt, Germany, pp.102-105, Oct. 15-17, 2010
- [8] Rajesh Mehta, Anurag Mishra, Rampal Singh and Navin Rajpal, “*Digital Image Watermarking in DCT Domain Using Finite Newton Support Vector Regression*”, In the Proceedings of Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Darmstadt, Germany, pp. 123 – 126, Oct. 15-17, 2010
- [9] Saraju P Mohanty, K R Ramakrishnan and Mohan Kankanhalli, “*A Dual Watermarking Technique for Images*”, ACM Multimedia, Part2, pp. 49-51, 1999
- [10] Mukesh C. Motwani and Fredrick C Harris, Jr, “*Fuzzy Perceptual Watermarking for Ownership Verification*”, in the Proc. 2009 International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV'09), Las Vegas, Nevada, 2009
- [11] Bassem Abdel-Aziz and Jean Yves Chouinard, “*On Perceptual Quality of Watermarked Images – An Experimental Approach*”, In Proc. 2nd International Workshop on Digital Watermarking (IWDW 2003), Seoul, Korea, pp. 277-288, Oct. 2003
- [12] A Shnayderman and A M. Eskicioglu, “*Evaluating the Visual Quality of Watermarked Images*”, IS&T/SPIE's 18th Annual Symposium on Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VIII Conference, San Jose, CA, Jan. 15–19, 2006
- [13] V V. Pankajakshan and F. Atrousseau, “*A Multi-Purpose Objective Quality Metric for Image Watermarking*”, in Proc. IEEE International Conference on Image Processing (ICIP 2010), pp. 2589-2592, Hong Kong, Sept. 26-29, 2010
- [14] Manish Narwaria, Weisi Lin, “*Scalable Image Quality Assessment Based on Structural Vectors*”, In Proc. IEEE International Conference on Multimedia Signal Processing MMSP '09, Rio de Janeiro, Brazil, Oct. 5-7, 2009
- [15] Cheng-Ri Piao, Seunghwa Beack, Dong-Min Woo, and Seung-Soo Han, “*A Blind Watermarking Algorithm Based on HVS and RBF Neural Network for Digital Image*”, In Proc. International Conference on Neural Computing (ICNC '06), Part I, LNCS 4221, pp. 493-496, 2006

Results: Original and Watermarked Images



FIGURE 4: Original Lena



FIGURE 5: Signed Image obtained by DCT scheme



FIGURE 6: Signed Image obtained by DWT scheme



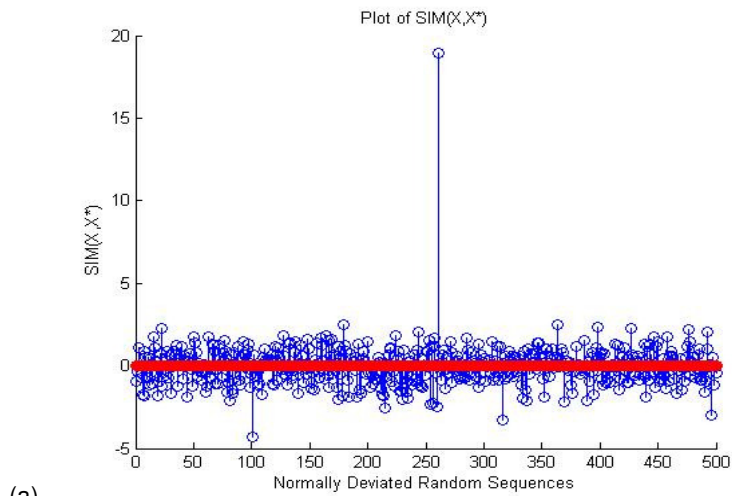
FIGURE 7: Signed Image obtained by RBF-NN scheme



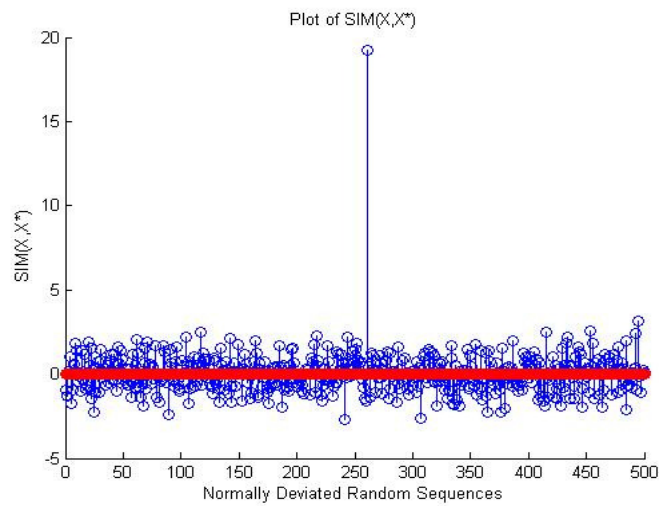
Figure 8(a-e): Images obtained after executing image processing attacks over signed image of Fig. 5 – (a) Counter clockwise rotation with 90° , (b) Gaussian Blur of radius = 1 unit, (c) Addition of Gaussian Noise of 10%, (d) Median Filtering (Filtering Aperture = 3) and (e) JPEG compression with quality factor $Q=0.9$



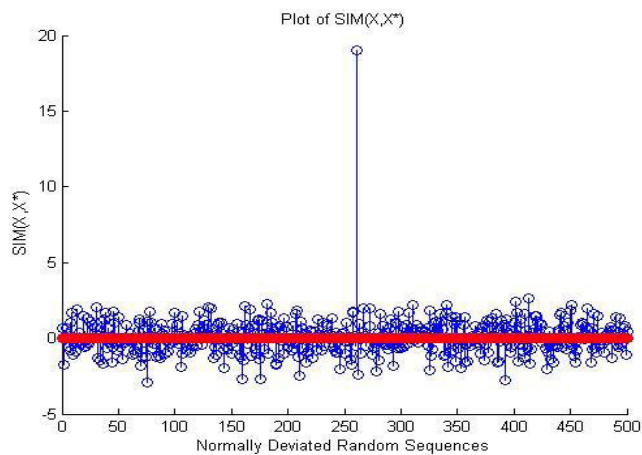




(a)



(b)



(c)

Fig. 11(a-c): $SIM(X, X^*)$ plots for random number watermark sequences extracted from images of Figs. 5-7 respectively. (a) DCT Watermarking scheme ($SIM = 18.9206$), (b) DWT Watermarking scheme ($SIM = 19.2357$) and (c) RBF-NN Watermarking scheme ($SIM = 19.0732$)

Fully Automatic Method for 3D T1-Weighted Brain Magnetic Resonance Images Segmentation

Bouchaib CHERRADI

*Ph. D student
UFR: MCM&SCP
F.S.T, BP 146, Mohammedia, Morocco.*

cherradi1@hotmail.com

Omar BOUATTANE

*Professor
E.N.S.E.T, Informatics Department
Hassan II St, Mohammedia, Morocco.*

bouattane@hotmail.com

Mohamed YOUSSEFI

*Ph. D student
F.S, Information Processing Department
Mohamed V University, Rabat. Morocco.*

med@youssefi.net

Abdelhadi RAIHANI

*Ph. D student
F.S, Information Processing Department
Hassan II University, Mohammedia. Morocco.*

abraihani@yahoo.fr

Abstract

Accurate segmentation of brain MR images is of interest for many brain disorders. However, due to several factors such noise, imaging artefacts, intrinsic tissue variation and partial volume effects, brain extraction and tissue segmentation remains a challenging task. So, in this paper, a full automatic method for segmentation of anatomical 3D brain MR images is proposed. The method consists of many steps. First, noise reduction by median filtering is done; second segmentation of brain/non-brain tissue is performed by using a Threshold Morphologic Brain Extraction method (TMBE). Then initial centroids estimation by gray level histogram analysis is executed, this stage yield to a Modified version of Fuzzy C-means Algorithm (MFCM) that is used for MRI tissue segmentation. Finally 3D visualisation of the three clusters (CSF, GM and WM) is performed. The efficiency of the proposed method is demonstrated by extensive segmentation experiments using simulated and real MR images. A confrontation of the method with similar methods of the literature has been undertaken through different performance measures. The MFCM for tissue segmentation introduce a gain in rapidity of convergence of about 70%.

Keywords: Noise Reduction, Brain Extraction, Clustering, MRI Segmentation, Performance Measures.

1. INTRODUCTION

Magnetic resonance (MR) imaging has been widely applied in biological research and diagnostics, primarily because of its excellent soft tissue contrast, non-invasive character, high spatial resolution and easy slice selection at any orientation. In many applications, its segmentation plays an important role on the following sides: (a) identifying anatomical areas of interest for diagnosis, treatment, or surgery planning paradigms; (b) pre-processing for multimodality image registration; and (c) improved correlation of anatomical areas of interest with localized functional metrics [1].

Intracranial segmentation commonly referred to as brain extraction, aims to segment the brain tissue (cortex and cerebellum) from the skull and non-brain intracranial tissues in magnetic

resonance (MR) images of the brain. Brain extraction is an important pre-processing step in neuroimaging analysis because brain images must typically be skull stripped before other processing algorithms such as registration, tissue classification or bias field correction can be applied [2-6]. In practice, brain extraction is widely used in neuroimaging analyses such as multi-modality image fusion and inter-subject image comparisons [2], [3]; examination of the progression of brain disorders such as Alzheimer's Disease [7, 8], multiple sclerosis [9-12] and schizophrenia [13], [14]; monitoring the development or aging of the brain [15], [16]; and creating probabilistic atlases from large groups of subjects [2]. Numerous automated brain extraction methods have been proposed [17-24]. However, the performance of these methods, which rely on signal intensity and signal contrast, may be influenced by numerous factors including MR signal inhomogeneities, type of MR image set, stability of system electronics, and extent of neurodegeneration in the subjects studied. In [25] we have proposed simple hybrid method, based on optimal thresholding and mathematical morphology operators for extracting brain tissues from 2D T1-weighted cerebral MRI images.

From the pattern recognition point of view, the tissue segmentation stage is to classify a set of elements defined by a set of features among which a set of classes can be previously known. In the MRI segmentation domain, the vector pattern X corresponds to the gray level of the studied point (pixel or voxel). From these approaches, one distinguishes the supervised methods where the class features are known a priori, and the unsupervised ones which use the features auto-learning. From this point of view, several algorithms have been proposed such as: c-means [26], fuzzy c-means (FCM) [27], adaptive fuzzy c-means [28], modified fuzzy c-means [29] using illumination patterns and fuzzy c-means combined with neutrosophic set [30].

Segmentation is a very large problem; it requires several algorithmic techniques and different computational models, which can be sequential or parallel using processor elements (PE), cellular automata or neural networks. In [31], we have presented Parallel implementation of c-means clustering algorithm to demonstrate the effectiveness and how the complexity of the parallel algorithm can be reduced in the reconfigurable mesh computer (RMC) computational model. In [32] the authors present the design, the modelling and the realisation of an emulator for this massively parallel re-configurable mesh computer.

Fully automatic brain tissue segmentation of magnetic resonance images (MRI) is of great importance for research and clinical study of much neurological pathology. The accurate segmentation of MR images into different tissue classes, especially gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF), is an important task. Moreover, regional volume calculations of these tissues may bring even more useful diagnostic information. Among them, the quantization of gray and white matter volumes may be of major interest in neurodegenerative disorders such as Alzheimer disease, in movements disorders such as Parkinson or Parkinson related syndrome, in white matter metabolic or inflammatory disease, in congenital brain malformations or prenatal brain damage, or in post traumatic syndrome. The automatic segmentation of brain MR images, however, remains a persistent problem. Automated and reliable tissue classification is further complicated by the overlap of MR intensities of different tissue classes and by the presence of a spatially smoothly varying intensity inhomogeneity.

In this paper we present fully automatic method for brain MRI volume segmentation. The system combines noise reduction by median filtering, the proposed TMBE method for non brain tissue removal, initial centroids estimation by gray level histogram analysis, and Fuzzy C-means Algorithm for tissue segmentation. Extensive experiments using simulated and real MR image data show that the proposed method can produce good segmentation results. Quantitative evaluation of the efficiency of the proposed method for brain extraction and tissue segmentation is confronted to some well known methods through standard performance measure in the literature.

The reminder of this paper is organized as follows. Section 2 presents the pre-processing procedure in which we represent our proposed method for brain extraction (TMBE) and a procedure for noise removing. Tissue classification method and performance measure are presented in section 3. Simulation results for the two main stages in the fully automatic method

for T1-weighted MRI images (Brain Extraction and tissue classification) are introduced in Section 4. Finally, conclusion and perspectives are given in section 5.

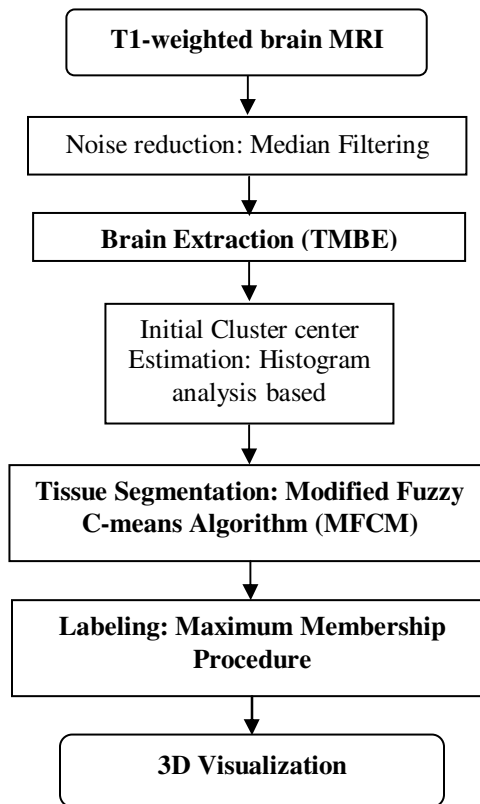


FIGURE 1: Global segmentation scheme.

2. PRE-PROCESSING

2.1. Noise reduction: Filtering

This pre-processing stage performs a non linear mapping of the grey level dynamics for the image. This transform consists in the application of a 3x3 median filter. The use of median filtering derives from the nature of the noise distribution in the MR images. The main source of noise in this kind of images is due to small density variations inside a single tissue which tend to locally modify the RF emission of the atomic nuclei during the imaging process. Such variations derive either from casual tissue motions or by external RF interferences and they assume a salt-and-pepper appearance. The median filter is a non-derivative low-pass one that removes efficiently this kind of disturb, allows homogeneous regions to become denser, thus improving clustering performance. The choice of the neighbourhood size derives from the need to avoid small regions to be confused with noise. Several approaches use complex models of the noise and perform anisotropic filtering because the noise distribution is, to some extent, oriented with the spatial direction of the RF atomic emission across the image. Here, the main concern is to reduce noise and make the single tissues more homogeneous.

2.2 Brain Extraction

The Brain extraction problem is a difficult task, which aims at extracting the brain from the skull, eliminating all non-brain tissue such as bones, eyes, skin, fat... This will allow us to simplify the segmentation of the brain tissues.

2.2.1 Some Previous Brain Extraction Techniques

2.2.1.1 Brain Extraction Tool (BET)

BET [21] is developed by FMRIB (Oxford Centre for Functional Magnetic Resonance Imaging of the Brain) and is available at <http://www.fmrib.ox.ac.uk/fsl/> for research purposes. In BET, the intensity histogram is processed to find "robust" lower and upper intensity values for the image, and a rough brain/non-brain threshold is determined. The center-of-gravity of the head image is found, along with the rough size of the head in the image. Next a triangular tessellation of a sphere's surface is initialized inside the brain, and allowed to slowly deform, one vertex at a time, following forces that keep the surface well-spaced and smooth, whilst attempting to move towards the brain's edge. If a suitably clean solution is not arrived at then this process is re-run with a higher smoothness constraint.

2.2.1.2 Brain Surface Extractor (BSE)

BSE [22] is developed by Neuroimaging Research Group, University of Southern California and the executable is available from <http://neuroimage.usc.edu/BSE/>. BSE is an edge based method that employs anisotropic diffusion filtering. Edge detection is implemented using a 2D Marr-Hildreth operator, employing low-pass filtering with a Gaussian kernel and localization of zero crossings in the Laplacian of the filtered image. The final step is morphological processing of the edge map.

2.2.1.3 McStrip (Minneapolis Consensus Stripping)

McStrip [16], [17] is developed by International Neuroimaging Consortium (INC) and is available for download at http://www.neurovia.umn.edu/incweb/McStrip_download.html. McStrip is initialized with a warp mask using AIR (<http://bishopw.loni.ucla.edu/AIR5/>), and dilates the AIR mask to form a Coarse Mask. It then estimates a brain/ non-brain threshold based on the intensity histogram, and automatically adjusts this threshold to produce a Threshold Mask. The volume of tissue within the Threshold Mask determines the choice of the BSE Mask from among a suite of 15 masks computed using parameter combinations spanning both smoothing and edge parameters. The final, McStrip Mask is a union of the Threshold and BSE masks after void filling and smoothing.

2.2.2 Threshold Morphologic Brain Extraction (TMBE)

Our simple and effective method is divided in five steps [25]

2.2.2.1 Binarisation by Thresholding

This step is based on global binary image thresholding using Otsu's method [33]. Figure 2-b shows a result of this operation.

2.2.2.2 Greatest Connected Component Extraction

A survey based on a statistical analysis of the existing connected components on the binary image, permits to extract the region whose area is the biggest. Figure 2-c shows a result of this operation.

2.2.2.3 Filling the Holes

The remaining holes in the binary image obtained in step 2, containing the greatest connected component, are filled using morphologic operation consisting of filling holes in the binary image. A hole is a set of background voxels within connected component. The result of this operation is shown in figure 2-d.

2.2.2.4 Dilatation

This morphologic operation consists of eliminating all remaining black spots on the white surface of the image. These spots are covered by the dilatation of the white parts. This carried out by moving a square structuring element of size $(S \cdot S)$ on binary image and applying logical OR operator on each of the $(S^2 - 1)$ neighbouring pixels (figure 2-e). Here we choose $S=3$.

2.2.2.5 Brain Extracting

The region of interest is the brain. To extract this region we use the AND operator between the original filtered image (figure 2-a) and the binary mask obtained in last step as is shown in figure 2-f. The non-brain tissues are obtained by applying AND operator between the image in figure 2a and the logical complement of the mask in figure 2e, the result is in figure 2-g.

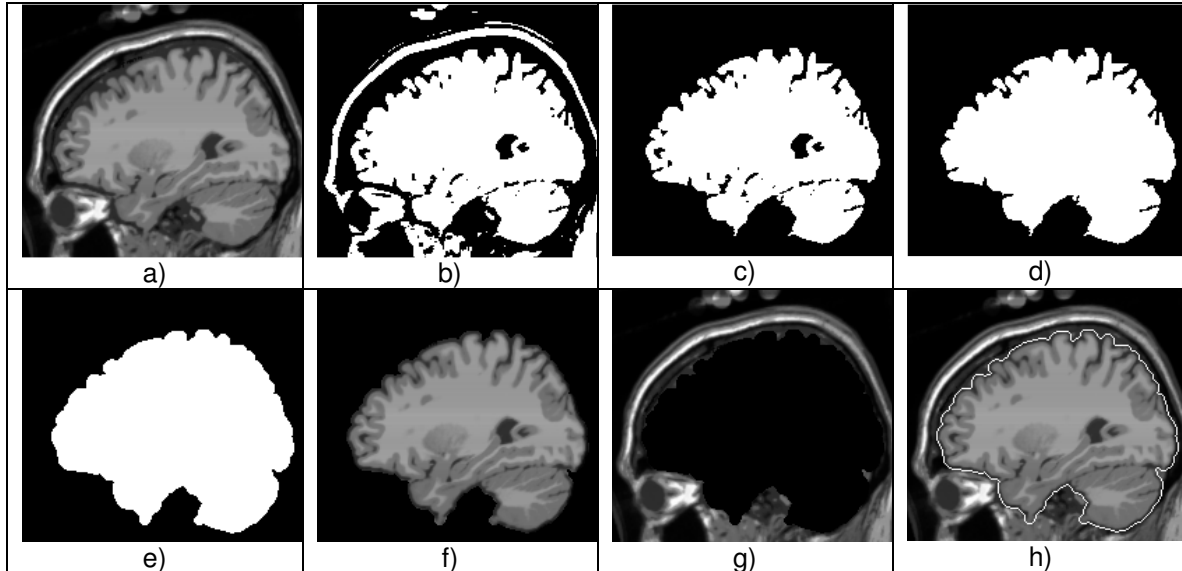


FIGURE 2: TMBE steps on sagittale slice number 120/181 from normal brain simulated phantom [40].

The figure 2-h shows the region of interest corresponding to the effective brain tissues in original MRI delimited by contour.

2.3. Histogram Based Centroids Initialization

Clustering algorithms requires an initialisation of the centroids values. Usually, this is randomly made. However, an adequate selection permits generally to improve the accuracy and reduces considerably the number of required iterations to the convergence of these algorithms. Among the methods used to estimate initial cluster values in the image, we used the histogram information analysis [25], [34].

The choice of the class number is done according to the quantity of information that we want to extract from the image. In our case this number is known in advance since we have extract three clusters from normal images (CSF, GM and WM).

2. TISSUE CLASSIFICATION

3.1. Image Segmentation

The objective of image segmentation is to divide an image into meaningful regions. Errors made at this stage would affect all higher level activities. In an ideally segmented image, each region should be homogeneous with respect to some criteria such as gray level, color or texture, and adjacent regions should have significantly different characteristics or features. More formally, segmentation is the process of partitioning the entire image into C regions $\{R_i\}$ such that each R_i is homogeneous with respect to some criteria. In many situations, it is not easy to determine if a voxel should belong to a region or not. This is because the features used to determine homogeneity may not have sharp transitions at region boundaries. To alleviate this situation, we inset fuzzy set concepts into the segmentation process. In fuzzy segmentation, each voxel is assigned a membership value in each of the C regions. If the memberships are taken into

account while computing properties of regions, we obtain more accurate estimates of region properties. One of the known techniques to obtain such a classification is the FCM algorithm.

3.2. Clustering: Modified FCM Algorithm

The fuzzy c-means (FCM) clustering algorithm was first introduced by DUNN [35] and later was extended by BEZDEK [36]. Fuzzy C-means (FCM) is a clustering technique that employs fuzzy partitioning such that a data point can belong to all classes with different membership grades between 0 and 1.

The aim of FCM is to find C cluster centers (centroids) in the data set $X = \{x_1, x_2, \dots, x_N\} \subseteq R_p$ that minimize the following dissimilarity function:

$$J_{FCM} = \sum_{i=1}^C J_i = \sum_{i=1}^C \sum_{j=1}^N u_{ij}^m d^2(V_i, x_j) \quad (1)$$

- u_{ij} : Membership of data x_j in the cluster V_i ;
- V_i : Centroid of cluster i ;
- $d_{(V_i, x_j)}$: Euclidian distance between i^{th} centroid (V_i) and j^{th} data point x_j ;
- $m \in [1, \infty[$: Fuzzy weighting exponent (generally equals 2).
- N : Number of data.
- C : Number of clusters, $2 \leq C \leq N$.
- p : Number of features in each data x_j .

With the constraints:

$$u_{ij} \in [0,1], \forall i, j \quad (2a)$$

$$\sum_{i=1}^C u_{ij} = 1, \forall j = 1, \dots, N \quad (2b)$$

$$0 < \sum_{j=1}^N u_{ij} < N, \forall i = 1, \dots, C \quad (2c)$$

To reach a minimum of dissimilarity function there are two conditions.

$$V_i = \frac{\sum_{j=1}^N u_{ij}^m x_j}{\sum_{j=1}^N u_{ij}^m} \quad (3)$$

$$u_{ij} = \frac{1}{\sum_{k=1}^C \left(\frac{d_{ij}}{d_{kj}} \right)^{2/(m-1)}} \quad (4)$$

We have modified this iterative algorithm to include the proposed procedure for estimating initial centroids with histogram analysis; this algorithm is in the following steps.

Step 0. Estimate the number of clusters C according to the procedure in section 2-3, choose the correspondent's gray level values as initial values of cluster centres $V^{(0)}$, Choose fuzzification parameter m ($1 < m < \infty$) $m=2$, and choose threshold $\epsilon > 0$. Initialize the membership matrix (U) according to the constraints of equations 2a, 2b and 2c.

At iteration N_i

{ **Step 1.** Calculate centroids $V^{(N_i)}$ using Equation (3).

Step 2. Compute dissimilarity function J_{N_i} using equation (1). If its improvement over previous iteration ($J_{N_i} - J_{N_i-1}$) is below a threshold $\epsilon > 0$, Go to Step 4.

Step 3. Compute a new membership matrix (U_{N_i}) using Equation (4). Go to Step 1.

Step 4. Stop. }

3.3. Performance Measures

To compare the performance of various segmentation techniques, we compute different coefficients reflecting how well two segmented regions match. The manually segmented regions are used as a gold standard (Truth Verity), and the automatically segmented ones are compared to them. To provide comparison between methods, we use a different performance measure:

3.3.1. Jaccard Similarity Coefficient

According to [37] the Jaccard similarity coefficient JSC is formulated as:

$$JSC = Card(R_1 \cap R_2) / Card(R_1 \cup R_2) \quad (5)$$

Where R_1 is the automatically segmented region, R_2 is the correspondent region of the manually segmented image, and $Card(X)$ denotes the number of voxels in the region X . A JSC of 1.0 represents perfect overlap, whereas an index of 0.0 represents no overlap. JSC values of 1.0 are desired.

3.3.2. Dice Similarity Coefficient [38]

Dice Similarity Coefficient is used to show the similarity level of automatically segmented region to manual segmented one. The Dice coefficient is defined as:

$$DSC = 2 * Card(R_1 \cap R_2) / Card(R_1 + R_2) \quad (6)$$

Where R_1 is the automatically segmented region, R_2 is the region of the manually segmented image, and $Card(X)$ denotes the number of voxels in the region X . A DSC of 1.0 represents perfect overlap, whereas an index of 0.0 represents no overlap. DSC values of 1.0 are desired.

3.3.3. Sensitivity and Specificity [39]

We also compute the sensitivity and specificity coefficient of the automated segmentation result using the manually segmented mask. The Sensitivity is the percentage of voxels recognized by the algorithm (Equation 7). The Specificity is the percentage of non recognized voxels by the algorithm (Equation 8).

$$Sensitivity = \frac{TP}{TP + FN} \quad (7)$$

$$Specificity = \frac{TN}{TN + FP} \quad (8)$$

Where TP and FP stand for true positive and false positive, which were defined as the number of voxels in R_1 correctly and incorrectly classified as R_2 by the automated algorithm. TN and FN stand for true negative and false negative, which were defined as the number of voxels in R_1 correctly and incorrectly classified as non R_2 by the automated algorithm.

4. RESULTS AND DISCUSSION

4.1. Brain Extraction

To prove the effectiveness of the proposed method for the skull stripping problem we have massively experiment TMBE using simulated and real MR image data in different modalities of acquisition. The figure 3 shows some samples of pre-processed images.

To evaluate the TMBE method we used a set of simulated and real volumes given from reference sites, they are presented as follows:

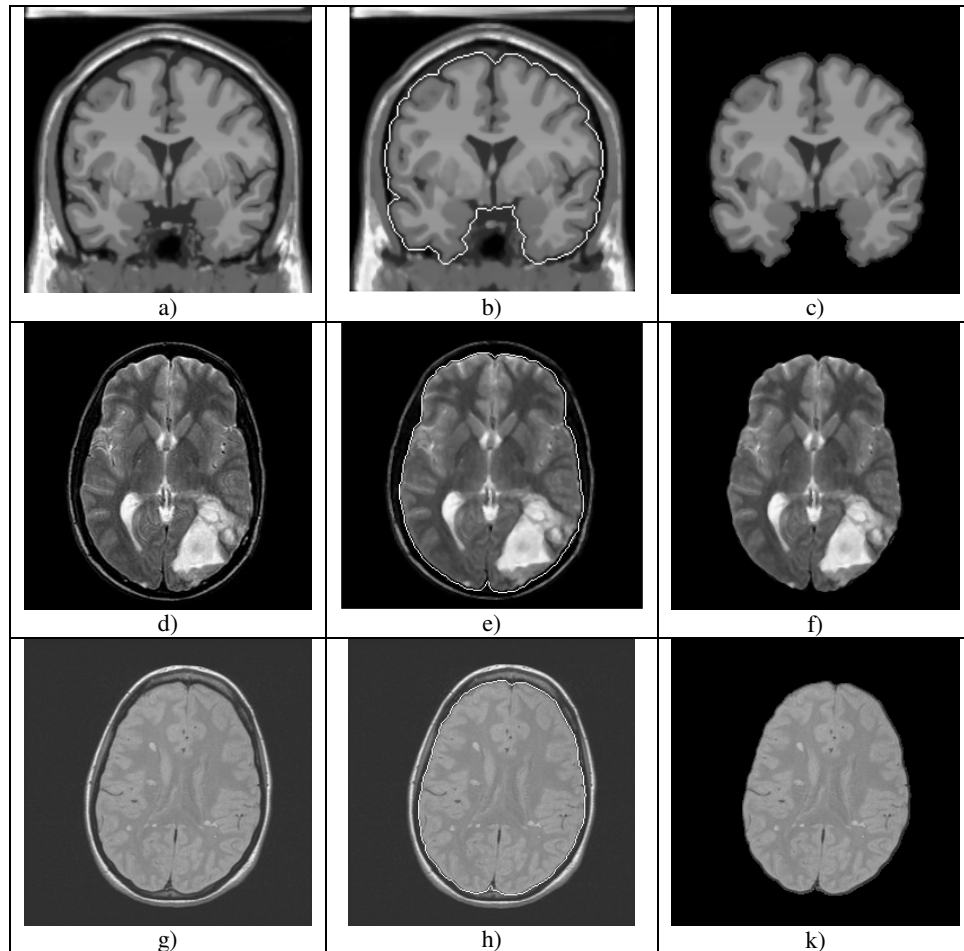


FIGURE 3: Some examples of pre-processed images by the proposed TMBE method (for Qualitative evaluation). a-c) Simulated T1-weighted image number 127/217 in coronal direction, d-f) T2-weighted image with tumor, g-k) PD-weighted image with Multiple Sclerosis (MS) lesion.

- 20 simulated volumes of size $181 \times 217 \times 181$ voxels given from the Brainweb simulated brain database [40] and their given manual segmentation (determined by union of the three tissues to form the correspondent region of interest (Brain) for each volume). This T1-weighted data are provided with $1mm \times 1mm \times 1mm$ in spacing.

- 18 real T1-weighted volumes which were acquired coronally with size $256 \times 256 \times 128$ voxels and $0.94mm \times 0.94mm \times 1.5mm$ as spatial resolution from the International Brain Segmentation Repository IBSR V2.0 [41]. The MR brain data sets and their manual segmentations in three tissues by expert radiologists were provided by the Center for Morphometric Analysis at Massachusetts General Hospital (The image data sets used were named IBSR_01 through IBSR_18).

T1-weighted modality, that belong to the fastest MRI modalities available, are often preferred, since they offer a good contrast between gray (GM) and white cerebral matter (WM) as well as between GM and cerebrospinal fluid (CSF).

To compare the performance of TMBE with three well known brain extraction techniques BET [21], BSE [22] and McStrip [23, 24] we compute the different coefficients described in section 3.3. The manually segmented brains are used as a Truth Verity (TV), and the automatically extracted brains by the proposed method TMBE are compared to them.

Quantitative comparison of the proposed Brain Extraction method TMBE to these brain extraction methods for two real datasets (IBSR_07 and IBSR_12) is summarised in Table.1 and Table.2. The values indicated in tables are average values for multiple essays.

Notice that, we have implemented the method in MATLAB 7.8 and have been used on a Pentium IV personal computer (Intel) with 2.6 GHz, 1024 MB of main memory, and an NVIDIA Geforce 7900 graphics card with 256 MB of graphics memory.

	JSC	DSC	Sensitivity	Specificity	time
BET [21]	0.81	0.76	0.603	0.912	3 min
BSE [22]	0.82	0.88	0.607	0.973	2 min
McStrip [23,24]	0.80	0.84	0.600	0.903	6 min
TMBE	0.80	0.87	0.599	0.901	4 min

TABLE 1: Different similarity index calculated for brain extraction of ISBR_7.

	JSC	DSC	Sensitivity	Specificity	time
BET [21]	0.85	0.78	0.600	0.902	3 min
BSE [22]	0.86	0.89	0.622	0.923	2 min
McStrip [23,24]	0.82	0.82	0.610	0.913	6 min
TMBE	0.84	0.86	0.591	0.923	4 min

TABLE 2: Different similarity index calculated for brain extraction of ISBR_12.

The comparison of TMBE with the three brain extraction techniques against expertly hand stripped T1-weighted MRI volumes revealed that TMBE method gives comparable results to BSE and BET in term of accuracy but with lowest time processing (creating mask in about 4 min). But when compared with McStrip technique it is faster.

4.2 Tissues Classification

The tissues classification aims to divide the extracted volume by TMBE in three clusters: Cerebrospinal fluid (CSF), gray matter (GM), and white matter (WM). The background voxels are removed by simple thresholding before the clustering starts.

4.2.1 Classification Results

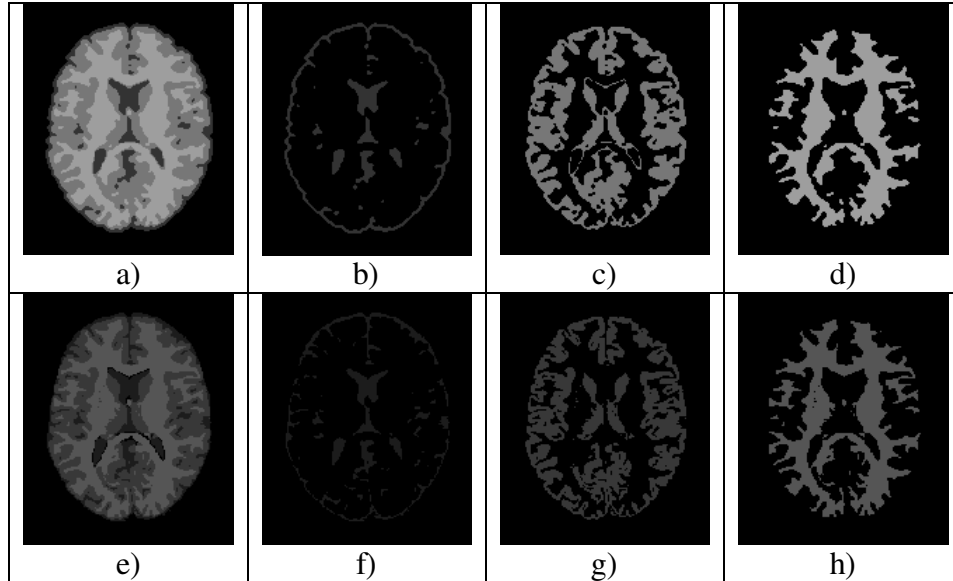


FIGURE 4: Example of segmentation results comparison. a) Segmented image by the proposed method, b) Cerebrospinal fluid (CSF) cluster, c) Gray matter (GM) cluster and d) The white matter (WM) cluster. e) Truth Verity image. f), g) and h) Manual segmentation of the same brain tissues (Brainweb).

For qualitative evaluation, Figure.4 shows segmentation results of axial T1-weighted slice of number 84 in axial direction obtained from the web site Brainweb [40] its about t1_icbm_normal_1mm_pn0_rf0 volume file which we call **Dataset1**, the image was segmented in three clusters (Truth Verity). It is very clear from this figure that the separation of the three clusters is very effective in comparison with the correspondent's results (TV).

4.2.2 Parameters Dynamic of the MFCM.

Table.3 shows the Different parameters states of the Modified FCM clustering by starting from centroids: (C1: CSF, C2: GM and C3: WM) = (53.50, 115.01 and 150.50), corresponding to the results in figure 4 and figure 5.

Ni	Value of each Cluster			Number of voxels in each Cluster			ObjFcn value
	CSF	GM	WM	Card(CSF)	Card(GM)	Card(WM)	J(Ni)
1	53.50	115.01	150.50	2364	7270	7617	1299906.70
2	52.93	113.42	151.71	2322	7312	7617	1293634.77
3	52.51	113.05	151.78	2288	7346	7617	1293020.71
4	52.31	112.92	151.76	2288	7346	7617	1292923.43
5	52.23	112.87	151.74	2288	7346	7617	1292907.25
6	52.20	112.86	151.74	2288	7346	7617	1292904.56

TABLE 3: Different parameters states of the clustering method starting from centroids: (C1: CSF, C2: GM and C3: WM) = (53.50, 115.01 and 150.50).

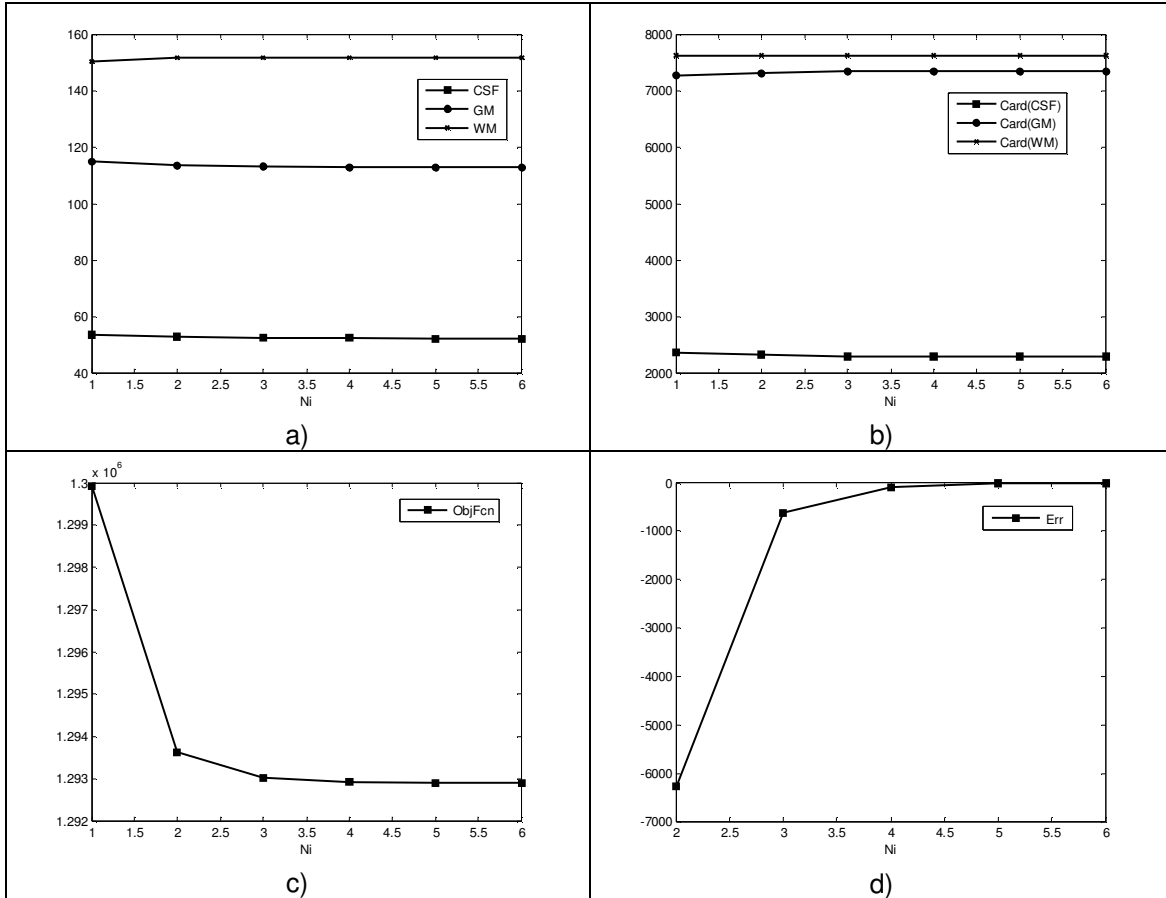


FIGURE 5: Dynamic of different clustering parameters:

- a) Centroids starting from (C1, C2, C3) = (35.50, 115.01, 150.50) as results of histogram analysis,
- b) Cardinality of each tissue, c) Value of objective function J(Ni), d) Value of Err=(J(Ni)-J(Ni-1)).

In figure.5 we present the dynamic of different clustering parameters using the results of histogram analysis leading to a centroids initialization of the extracted region of interest consisting of brain tissues that we want segment. These results correspond to variation trough the iterations of the clusters centroids values, the cardinality of each cluster, the FCM objective function J(Ni) and the difference between two successive objective function values (J(Ni)-J(Ni-1)) used as criteria for convergence.

As shown in figure. 5 and trough many experiment done on different images, the rapidity of the method is much enhanced (6 iterations) in comparison with the case of random initialisation of the cluster centroids that is practiced in standard FCM clustering (about 20 iterations).

Table.3 shows the dynamic of different parameters states of the modified FCM clustering starting from centroids: (C1: CSF, C2: GM and C3: WM) = (53.50, 115.01 and 150.50), corresponding to the results presented in figures 4 and 5, this values of initial centroids are obtained by histogram analysis described in section 2-3.

In this table we show that we can stop iterative procedure more early, since the clusters cardinalities don't change anymore from the iteration number 3. The clusters centres have their lower change with very small amounts.

4.2.3 Noise Robustness of the Proposed Method

To evaluate the robustness of the proposed method for tissue classification to the presence of noise, we used **Dataset1** with additional noise levels: 0%, 1%, 3%, 5%, 7% and 9%.

Dataset1: t1_icbm_normal_1mm_pn0_rf0 [40], is simulated normal brain phantom of 181x217x181 voxels with 1mm³ for each voxel without any noise or intensities inhomogeneity.

Noise	JSC	DSC	Sensitivity	Specificity
0%	0.942	0.955	0.757	0.895
1%	0.948	0.953	0.743	0.895
3%	0.919	0.943	0.732	0.899
5%	0.893	0.929	0.715	0.899
7%	0.854	0.922	0.704	0.897
9%	0.849	0.908	0.688	0.897

TABLE 4: Performance measures for Modified FCM Clustering results of WM tissue (Dataset1).

In table.4 and table.5 we summarise the performance measure results calculated for segmented GM and WM of different variants of **Dataset1** obtained with additional noise of different amounts.

Noise	JSC	DSC	Sensitivity	Specificity
0%	0.952	0.975	0.657	0.897
1%	0.948	0.973	0.653	0.887
3%	0.929	0.963	0.642	0.880
5%	0.903	0.949	0.625	0.870
7%	0.874	0.932	0.614	0.850
9%	0.849	0.918	0.598	0.840

TABLE 5: Performance measures for Modified FCM Clustering results of GM tissue (Dataset1).

4.2.4 Visualisation of 3D Rendered Surface.

To appreciate the segmentation results obtained slice by slice, in 3D space, we export our segmentation results to ANALYZE 10.0 that is a comprehensive and interactive package for multidimensional image visualization, processing and analysis developed by The Biomedical Imaging Resource at Mayo Clinic, Rochester, MN [42].

The figure.6 shows 3D rendered surface of the segmentation results for the three tissues extracted from Dataset1. For the CSF we have limited the visualisation to the lateral ventricles.

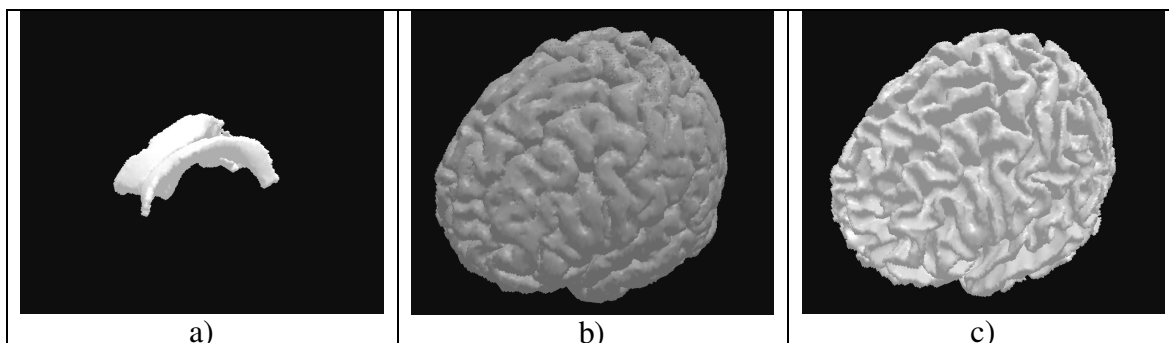


FIGURE 6: 3D visualization of rendered surface for segmented volume (Dataset1). a) Lateral Ventricular CSF, b) GM, c) WM.

5. CONCLUSION AND PERSPECTIVES

In this paper, we have presented a complete MRI images segmentation method. Unlike other brain segmentation methods described in the literature, the one described in this paper is truly automatic because it does not require a user to determine image-specific parameters, thresholds, or regions of interest.

The automatic proposed method for extracting the brain from the T1-weighted MRI head scans is based on a hybrid processing techniques including global optimal thresholding and mathematical morphology operators. Our quantitative results show that the proposed method achieves comparable performance with synthetic BrainWeb data, and real IBSR V2.0 data against standard techniques such as BSE and BET, and **McStrip**. Our results are also more consistent across the datasets, making the proposed method suited for measuring brain volumes in a clinical setting.

Concerning the tissues classification we used modified FCM that is a clustering technique that utilizes the distance between voxels and cluster centres in the spatial domain to compute the membership function. The modification consists of using the histogram analysis for the determination of initial cluster centroids instead of a random initialization. The segmentation process is achieved in 6 iterations instead of about 20 iterations when we used standard FCM with random initial centroids. This is important improvement (about 70%) especially in our case where we manipulate big quantity of data.

The accuracy and the effectiveness of the fully automatic proposed method for 3D brain MR images segmentation has been evaluated qualitatively and quantitatively, but more work can be done to improve the method that need to be tested on many more data sets to expose unexpected segmentation errors that might occur infrequently. The method should also be tested with more recent images database.

More comprehensive comparison of MFCM with other clustering models will be addressed. Future work will focus on developing an automatic image based classification system for brain tumor using MRI data of different modalities and taking into account the intensity nonuniformity artefact.

Another improvement in time processing can be gained while modifying the convergence criteria for FCM clustering by considering a threshold on the change of clusters centres instead of calculating the objective function.

An implementation of the proposed method as well as other algorithms for MRI segmentation on massively parallel reconfigurable mesh computer emulator [32] is being finalised.

REFERENCES.

- [1] L.P Clarke, R.P Velthijzen, M.A Camacho, J. J Heine, "MRI segmentation methods and applications", Magnetic Resonance Imaging, Vol. 13, No. 3, pp. 343-368, 1995.
- [2] R. P. Woods, S. T. Grafton, J. D. G. Watson, N. L.Sicotte, and J. C. Mazziotta, "Automated image registration: II. Intersubject validation of linear and nonlinear models", J. of Computer Assisted Tomography, vol. 22, pp: 139-152, 1998.
- [3] R. P. Woods, M. Dapretto, N. L. Sicotte, A. W. Toga, and J. C. Mazziotta, "Creation and use of a Talairach-Compatible atlas for accurate, automated, nonlinear intersubject registration, and analysis of functional imaging data", Human Brain Mapping, vol. 8, pp: 73-79, 1999.
- [4] J. Van Horn, T. M. Ellmore, G. Esposito, K. F. and Berman, "Mapping Voxel-based Statistical Power on Parametric Imaging", NeuroImage, vol. 7, pp: 97-107, 1998.

- [5] D. W. Shattuck, S.R. Sandor-Leahy, K. A. Schaper, D. A. Rottenberg, and R. M. Leahy, "Magnetic Resonance Image Tissue Classification Using a Partial Volume Model", *NeuroImage*, vol. 13, pp: 856-876, 2001.
- [6] S. Strother, S. La Conte, L. Kai Hansen, J. Anderson, J. Zhang, S. Pulapura, and D. Rottenberg, "Optimizing the fMRI data-processing pipeline using prediction and reproducibility performance metrics: I. A preliminary group analysis", *NeuroImage*, vol. 23, pp:196-207, 2004.
- [7] H. Rusinek, M. J. de Leon, A. E. George, L. A. Stylopoulos, R. Chandra, G. Smith, T. Rand, M. Mourino, and H. Kowalski, "Alzheimer disease: measuring loss of cerebral gray matter with MR imaging", *Radiology*, vol. 178, pp: 109-114, 1991.
- [8] P. M Thompson, M. S. Mega, R. P. Woods, C. I. Zoumalan, C. J. Lindshield, R. E. Blanton, J. Moussai, C. J. Holmes, J. L. Cummings, and A. W. Toga, "Cortical change in Alzheimer's disease in detected with a disease specific population-based brain atlas", *Cerebral Cortex* vol. 11, pp: 1-16, 2001.
- [9] R. A. Bermel, J. Sharma, C. W. Tjoa, S. R. Puli, and R. Bakshi, "A semiautomated measure of whole- brain atrophy in multiple sclerosis", *Neurological Sciences*, vol. 208, pp: 57-65, 2003.
- [10] M. A. Horsfield, M. Rovaris, M. A. Rocca, P. Rossi, R. H. B. Benedict, M. Filippi, and R. Bakshi, "Whole-brain atrophy in multiple sclerosis measured by two segmentation processes from various MRI sequences", *Neurological Sciences*, vol. 216, pp: 169-177, 2003.
- [11] R. Zivadinov, F. Bagnato, D. Nasuelli, S. Bastianello, A. Bratina, L. Locatelli, K. Watts, L. Finamore, A. Grop, M. Dwyer, M. Catalan, A. Clemenzi, E. Millefiorini, R. Bakshi, and M. Zorzon, M., " Short-term brain atrophy changes in relapsing-remitting multiple sclerosis", *Neurological Sciences*, vol. 223, pp: 185-193, 2004.
- [12] J. Sharma, M. P. Sanfilippo, R. H. B. Benedict, B. Weinstock-Guttman, F. E. Munschauer, and R. Bakshi, "Whole-brain atrophy in multiple sclerosis measured by automated versus semi-automated MR imaging segmentation", *Neuroradiology*, vol. 25, pp: 985-996, 2004.
- [13] K. L. Narr, P. M. Thompson, P. Szeszko, D. Robinson, S. Jang, R. P. Woods, S. Kim, K. M. Hayashi, D. Asuncion, A. W. Toga, and R. M. Bilder, "Regional specificity of hippocampal volume reductions in first episode schizophrenia", *NeuroImage*, vol. 21, pp: 1563- 1575, 2004.
- [14] P. Tanskanen, J. M. Veijola, U. K. Piippo, M. Haapea, J. A. Miettunen, J. Pyhtinen, E. T. Bullmore, P. B. Jones, and M. K. Isohanni, "Hippocampus and amygdale volumes in schizophrenia and other psychoses in the Northern Finland 1966 birth cohort". *Schizophrenia Research*, vol. 75, pp: 283-294, 2005.
- [15] T. L. Jernigan, S. L. Archibald, C. Fennema-Notestine, A. C. Gamst, J. C. Stout, J. Bonner, and J. R. Hesselink, "Effects of age on tissues and regions of the cerebrum and cerebellum", *Neurobiology of Aging*, vol. 22, pp: 581-594, 2001.
- [16] R. E. Blanton, J. G. Levitt, J. R. Peterson, D. Fadale, M.L. Sporty, M. Lee, D. To, E. C. Mormino, P. M. Thompson, J. T. McCracken, and A. W. Toga, "Gender differences in the left inferior frontal gyrus in normal children", *NeuroImage*, vol. 22, pp: 626-636, 2004.
- [17] A.M. Dale AM, B. Fischl , and M. Sereno, "Cortical surface-based analysis. Segmentation and surface reconstruction", *Neuroimage*, vol. 9, pp: 179-94, 1999.

- [18] H. Hahn, and H-O. Peitgen, "The skull stripping problem in MRI solved by a single 3D watershed transform", MICCAI, LNCS 1935, pp: 134-143, 2000.
- [19] S. Sandor, and R. Leahy, "Surface-based labeling of cortical anatomy using a deformable database", IEEE Transactions on Medical Imaging, vol. 16, pp: 41-54, 1997.
- [20] F. Segonne, A. M. Dale, E. Busa, M. Glessner, D. Salat, H. K. Hahn, and B. Fischl, "A hybrid approach to the skull stripping problem in MRI", Neuroimage, vol. 22, pp :1060-75, 2004.
- [21] S. M. Smith, "Fast robust automated brain extraction", Human Brain Mapping, vol. 17, pp: 143-55, 2002.
- [22] Shattuck, D.W., Sandor-Leahy, S.R., Shaper, K.A., Rottenberg, D.A., Leahy, R.M., Magnetic resonance image tissue classification using a partial volume model. NeuroImage. 13 (5), 856–876. 2001.
- [23] Rehm K, Shattuck D, Leahy R, Schaper K, Rottenberg DA. "Semi-automated stripping of T1 MRI volumes: I. Consensus of intensity- and edge-based methods". NeuroImage. 9(6) S86, 1999.
- [24] Rehm K, Schaper K, Anderson J, Woods R, Stoltzner S, Rottenberg D. "Putting our heads together: a consensus approach to brain/non-brain segmentation in T1-weighted MR volumes". NeuroImage, November, 2003.
- [25] B. Cherradi O. Bouattane, M. Youssfi and A. Raihani," Brain Extraction and Fuzzy Tissue Segmentation in Cerebral 2D T1-Weighted Magnetic Resonance Images". International Journal of Computer Science Issues, Vol. 8, Issue 3, May 2011, In Press
- [26] J. B. MacQueen, "Some Methods for classification and Analysis of Multivariate Observations", Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, University of California Press, 1:281-297,1967.
- [27] J.S.R. Jang, C. Sun, T.E. Mizutani, Neuro-Fuzzy and Soft Computing, Prentice Hall, pp. 426–427. 1997.
- [28] L. Dzung Pham, L.P. Jerry, "An adaptative fuzzy c-means algorithm for image segmentation in the presence of intensity inhomogeneities", Pattern Recognition Letters 20 57–68,1999.
- [29] L. Ma, R.C. Staunton, "A modified fuzzy c-means image segmentation algorithm for use with uneven illumination patterns", Pattern Recognition 40 3005–3011, 2007.
- [30] Y. Guo, H.D. Cheng, W. Zaho, Y. Zhang, "A novel image segmentation algorithm based on fuzzy c-means algorithm and neutrosophic set", in: Proceeding of the 11th Joint Conference on Information Sciences, Atlantis Press, 2008.
- [31] O. Bouattane, B. Cherradi, M. Youssfi and M.O. Bensalah. "Parallel c-means algorithm for image segmentation on a reconfigurable mesh computer". Parallel Computing, Volume 37, Issues 4-5, Pages 230-243, April-May 2011.
- [32] M. Youssfi, O. Bouattane, M.O. Bensalah, "A massively parallel re-configurable mesh computer emulator: design, modeling and realization", Journal of Software Engineering and Applications vol 3, pp 11–26, 2010.
- [33] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms". IEEE Transactions on Systems, Man and Cybernetics, Vol. 9, No. 1, pp. 62-661979.
- [34] B. Cherradi, O.Bouattane, M. Youssfi and A. Raihani, "Fuzzy segmentation method for MRI images with spatial information", in the proceeding of Mediterranien Congres of Telecommunications (CMT2010). Casablanca, Morocco. pp 296-299, 2010.

- [35] J.C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters", *Journal of Cybernetics* 3(3), pp. 32–57, 1973.
- [36] J.C. Bezdek, "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum Press, New York 1981.
- [37] P. Jaccard, "The distribution of the flora in the alpine zone". *New Phytol.* 11 (2), 37–50, 1912.
- [38] L. Dice, "Measures of the amount of ecologic association between species", *Ecology*, vol. 26, pp: 297- 302, 1945.
- [39] MJ. Gardner and DG. Altman, "Calculating confidence intervals for proportions and their differences". BMJ Publishing Group, pp 28-33, 1989.
- [40] <http://www.bic.mni.mcgill.ca/brainweb/>.
- [41] <http://www.cma.mgh.harvard.edu/ibsr/>.
- [42] <http://www.mayo.edu/bir/Software/Analyze/Analyze.html>.

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Image Processing (IJIP)* aims to be an effective forum for interchange of high quality theoretical and applied research in the Image Processing domain from basic research to application development. It emphasizes on efficient and effective image technologies, and provides a central forum for a deeper understanding in the discipline by encouraging the quantitative comparison and performance evaluation of the emerging components of image processing.

We welcome scientists, researchers, engineers and vendors from different disciplines to exchange ideas, identify problems, investigate relevant issues, share common interests, explore new approaches, and initiate possible collaborative research and system development.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for IJIP.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 5, 2011, IJIP appears in more focused issues. Besides normal publications, IJIP intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

LIST OF TOPICS

The realm of International Journal of Image Processing (IJIP) extends, but not limited, to the following:

- Architecture of imaging and vision systems
- Character and handwritten text recognition
- Chemistry of photosensitive materials
- Coding and transmission
- Color imaging
- Data fusion from multiple sensor inputs
- Document image understanding
- Holography
- Image capturing, databases
- Image processing applications
- Image representation, sensing
- Implementation and architectures
- Materials for electro-photography
- New visual services over ATM/packet network
- Object modeling and knowledge acquisition
- Photographic emulsions
- Prepress and printing technologies
- Remote image sensing
- Autonomous vehicles
- Chemical and spectral sensitization
- Coating technologies
- Cognitive aspects of image understanding
- Communication of visual data
- Display and printing
- Generation and display
- Image analysis and interpretation
- Image generation, manipulation, permanence
- Image processing: coding analysis and recognition
- Imaging systems and image scanning
- Latent image
- Network architecture for real-time video transport
- Non-impact printing technologies
- Photoconductors
- Photopolymers
- Protocols for packet video
- Retrieval and multimedia

- Storage and transmission

- Video coding algorithms and technologies for ATM/p

CALL FOR PAPERS

Volume: 5 - Issue: 4 - July 2011

i. Paper Submission: July 31, 2011

ii. Author Notification: September 01, 2011

iii. Issue Publication: September / October 2011

CONTACT INFORMATION

Computer Science Journals Sdn Bhd

M-3-19, Plaza Damas Sri Hartamas
50480, Kuala Lumpur MALAYSIA

Phone: 006 03 6207 1607
006 03 2782 6991

Fax: 006 03 6207 1697

Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2011
COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6207 1607
006 03 2782 6991

FAX: 006 03 6207 1697
EMAIL: cscpress@cscjournals.org

CSC PUBLISHERS © 2011
COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6207 1607
006 03 2782 6991

FAX: 006 03 6207 1697
EMAIL: cscpress@cscjournals.org