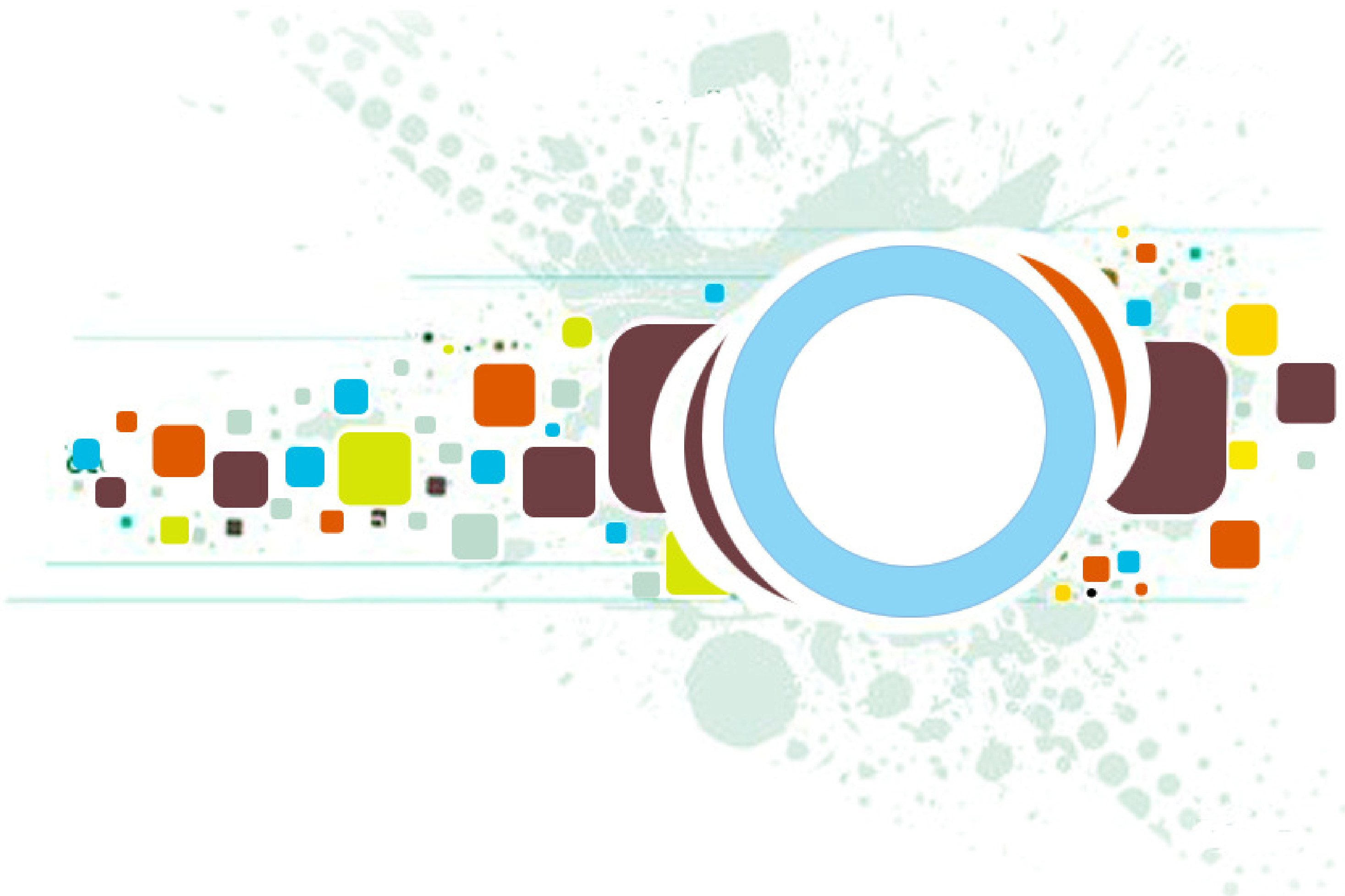
Volume 5 ▪ Issue 1 ▪ March 2011

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 1, 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 1, April 2011
Publishing Date: 04-04-2011
ISSN (Online): 1985-2304

IJIP Journal is a part of CSC Publishers
http://www.cscjournals.org

Published in Malaysia

Typesetting: Camera-ready by author, data conversation 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)

TABLE OF CONTENTS

Volume 5, Issue 1, December 2011

Pages

Content Based Image Retrieval Using Full Haar Sectorization

H.B.Kekre hbkekre@yahoo.com
Sr. Professor, Computer Engineering Dept
SVKM's NMIMS (Deemed-to be University)
Mumbai, 400056, INDIA

Dhirendra Mishra dhirendra.mishra@gmail.com
Associate Professor & PhD Research Scholar
SVKM's NMIMS (Deemed-to be University)
Mumbai, 400056, INDIA

Abstract

Content based image retrieval (CBIR) deals with retrieval of relevant images from the large image database. It works on the features of images extracted. In this paper we are using very innovative idea of sectorization of Full Haar Wavelet transformed images for extracting the features into 4, 8, 12 and 16 sectors. The paper proposes two planes to be sectored i.e. Forward plane (Even plane) and backward plane (Odd plane). Similarity measure is also very essential part of CBIR which lets one to find the closeness of the query image with the database images. We have used two similarity measures namely Euclidean distance (ED) and sum of absolute difference (AD). The overall performance of retrieval of the algorithm has been measured by average precision and recall cross over point and LIRS, LSRR. The paper compares the performance of the methods with respect to type of planes, number of sectors, types of similarity measures and values of LIRS and LSRR.

Keywords: CBIR, Haar Wavelet, Euclidian Distance, Sum of Absolute Difference, LIRS, LSRR, Precision and Recall.

1. INTRODUCTION

Content based image retrieval i.e. CBIR [1-4] is well known technology being used for the retrieval of images from the large database. CBIR has been proved to be very much needed technology to be researched on due to its applicability in various applications like face recognition, finger print recognition, pattern matching [15][17][21], verification /validation of images etc. The concept of CBIR can be easily understood by the figure 1 as shown below. Every CBIR systems needs to have module to extract features of an image it could be shape, color, feature which can be used as unique identity of the image. The features of the query image are compared with the features of all images in the large database using various similarity measures. These mathematical similarity measuring techniques checks the similarity of features extracted to classify the images in the relevant and irrelevant classes. The research in CBIR needs to be done to explore two things first is the better method of feature extraction having maximum components of uniqueness and faster the accurate mathematical models of similarity measures.

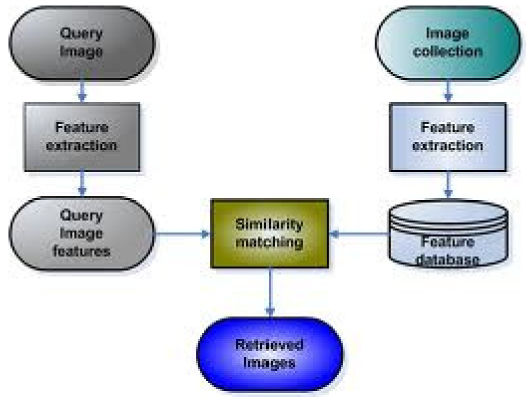


FIGURE 1: The CBIR System

There are lots of researches going on in the field of CBIR to generate the better methodologies of feature extractions in both spatial domain and frequency domain .Some methodologies like block truncation coding[19-21], clustering like vector quantization[18], various transforms: Kekre's Wavelet[5], DCT[16], DST[21][24], FFT[6-9], Walsh[10-11][22-23], Contourlet transform[3], using various methodologies like Complex Walsh sectors [12-14][17-23] has already been developed.

In this paper we have introduced a novel concept of Sectorization of Full Haar Wavelet transformed color images for feature extraction. Two different similarity measures parameters namely sum of absolute difference and Euclidean distance are considered. Average precision, Recall, LIRS and LSRR are used for performances study of these approaches.

2. HAAR WAVELET [5]

The Haar transform is derived from the Haar matrix. The Haar transform is separable and can be expressed in matrix form

$$[F] = [H]\ [f]\ [H]T$$

Where f is an NxN image, H is an NxN Haar transform matrix and F is the resulting NxN transformed image. The transformation H contains the Haar basis function hk(t) which are defined over the continuous closed interval t Є [0,1].

The Haar basis functions are When k=0, the Haar function is defined as a constant

$$hk(t) = 1/\sqrt{N}$$

When k>0, the Haar Function is defined as

$$h_k(t) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2} & (q-1)/2^p <= t < (q-0.5)/2^p \\ -2^{p/2} & (q-0.5)/2^p <= t < q/2^p \\ 0 \text{ otherwise} & \end{cases} \qquad (1)$$

$$\text{Where } 0 \le p < \log_2 N \text{ and } 1 \le q \le 2^p$$

3. SECTORIZATION OF TRANSFORMED IMAGES [8-14]

3.1 Haar Plane Formation

The components of Full Haar transformed image shown in the red bordered area (see Figure 2) are considered for feature vector generation. The average value of zeoeth row, column and last row and column components are considered only for augmentation purpose. We have used color codes to differentiate between the co-efficients plotted on Forward (Even) plane as light red and light blue for co-efficients belonging to backward (Odd) plane. The co-efficient with light red background i.e. at position (1,1),(2,2);(1,3),(2,4) etc. are taken as X1 and Y1 respectively and plotted on Even plane. The co-efficient with light blue background i.e. at position (2,1),(1,2);(2,3),(1,4) etc. are taken as X2 and Y2 respectively and plotted on Odd plane.

	0	1	2	3	4	5
0	0,0	0,1	0,2	0,3	0,4	0,5
1	1,0	1,1	1,2	1,3	1,4	1,5
2	2,0	2,1	2,2	2,3	2,4	2,5
3	3,0	3,1	3,2	3,3	3,4	3,5
4	4,0	4,1	4,2	4,3	4,4	4,5
5	5,0	5,1	5,2	5,3	5,4	5,5

FIGURE 2: Haar component arrangement in an Transformed Image.

Even plane of Full Haar is generated with taking Haar components into consideration as all X(i,j), Y(i+1, j+1) components for even plane and all X(i+1, j), Y(i, j+1) components for odd plane as shown in the Figure 3. Henceforth for our convenience we will refer X(i,j) = X1, Y(i+1,j+1) =Y1 and X(i+1,j) = X2 and Y(i,j+1) = Y2.

X(i,j)	Y(i,j+1)
X(i+1,j)	Y(i+1,j+1)

FIGURE 3: Snapshot of Components considered for Even/Odd Planes.

As shown in the Figure 3 the Even plane of Full Haar considers X1 i.e. all light red background cells (1,1), (2,2),(1,3),(2,4) etc. on X axis and Y1 i.e. (1,2), (2,1),(1,4),(2,3) etc. on Y axis. The Odd plane of Full Haar considers X1 i.e. all light blue background cells (1,2), (2,1),(1,4),(2,3) etc. on X axis and Y1 i.e. (1,2), (2,1),(1,4),(2,3) etc. on Y axis.

3.2 4 Sector Formation

Even and odd rows/columns of the transformed images are checked for sign changes and the based on which four sectors are formed as shown in the Figure 4 below:

Sign of X1/X2	Sign of Y1/Y2	Quadrants
+	+	I (0 – 900)
+	-	II (90 – 1800)
-	-	III(180- 2700)
-	+	IV(270–3600)

FIGURE 4: Computation of 4 Sectors

3.3. 8 Sectors Formation

The transformed image sectored in 4 sectors is taken into consideration for dividing it into 8 sectors. Each sector is of angle 45o. Coefficients of the transformed image lying in the particular sector checked for the sectorization conditions as shown in the Figure 5.

Sectors	Conditions
I,IV,V,VIII	$\|A\| >= \|B\|$
II,III,VI,VII	$\|B\| >= \|A\|$
Where A = X1 / X2 Component of the Transformed Image B = Y1 / Y2 Component of the Transformed Image	

FIGURE 5: Computation of 8 Sectors

3.4 12 Sector Formation

Division each sector of 4 sectors into angle of 30^o forms 12 sectors of the transformed image. Coefficients of the transformed image are divided into various sectors based on the inequalities shown in the Figure 6.

Sectors	Conditions
I, IV, VII, X	$\|A\| >= \sqrt{3} * \|B\|$
II, V, VIII, XI	$1/\sqrt{3} * \|A\| <= \|B\| <= \sqrt{3} * \|A\|$
III,VI, IX, XII	Otherwise
Where A = X1 / X2 Component of the Transformed Image B = Y1 / Y2 Component of the Transformed Image	

FIGURE 6: Computation of 12 Sectors

3.5 16 Sector Formation

Similarly we have done the calculation of inequalities to form the 16 sectors of the transformed image. The even/odd rows/ columns are assigned to particular sectors for feature vector generation

4. RESULTS OF EXPERIMENTS

We have used the augmented Wang image database [2] The Image database consists of 1055 images of 12 different classes such as Flower, Sunset, Barbie, Tribal, Cartoon, Elephant, Dinosaur, Bus, Scenery, Monuments, Horses, Beach. Class wise distribution of all images in the database has been shown in the Figure 7.

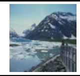
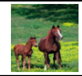
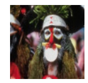
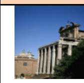
Class						
No. of Images	45	59	51	100	100	100
Class						
No. of Images	100	100	100	100	100	100

FIGURE 7: Class wise distribution of images in the Image database

FIGURE8: Query Image

The query image of the class dinosaur has been shown in Figure 8. For this query image the result of retrieval of both approaches of Full Haar wavelet transformed image sectorization of even and odd planes. The Figure 9 shows First 20 Retrieved Images sectorization of Full Haar wavelet Forward (Even) plane (16 Sectors) with sum of absolute difference as similarity measure. It can be observed that the retrieval of first 20 images are of relevant class i.e. dinosaur; there are no irrelevant images till first 77 retrievals in first case. The result of odd plane sectorization shown in Figure 10; the retrieval of first 20 images has 2 irrelevant images and 18 of relevant class.

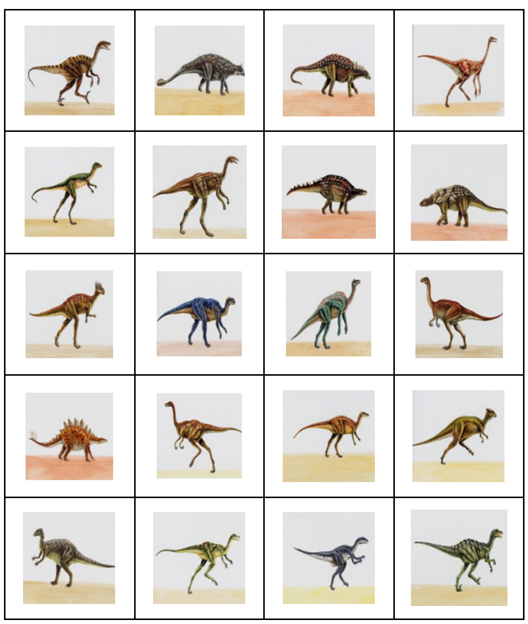


FIGURE 9: First 20 Retrieved Images sectorization of Full Haar wavelet Forward (Even) plane(16 Sectors)

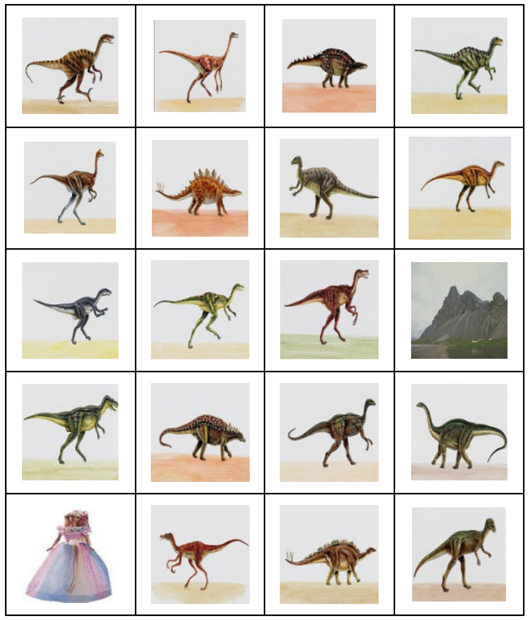


FIGURE 10: First 20 Retrieved Images of Full Haar wavelet Backward (odd) plane Sectorization (16 Sectors).

Feature database includes feature vectors of all images in the database. 5 randomly chosen query images of each class produced to search the database. The image with exact match gives minimum sum of absolute difference and Euclidian distance. To check the effectiveness of the work and its performance with respect to retrieval of the images we have calculated the overall average precision and recall as given in Equations (2) and (3) below. Two new parameters i.e. LIRS and LSRR are introduced as shown in Equations (4) and (5).

$$\text{Precision} = \frac{\text{Number of Relevant Images Retrieved}}{\text{Total Number of Images Retrieved}} \quad (2)$$

$$\text{Recall} = \frac{\text{Number of Relevant Images Retrieved}}{\text{Total Number of Relevant Images in the Database}} \quad (3)$$

$$\text{LIRS} = \frac{\text{Length of Initial Relevant string of images}}{\text{Total Relevant images Retrieved}} \quad (4)$$

$$\text{LSRR} = \frac{\text{Length of string to Recover all Relevant Images}}{\text{Total Images in the Database}} \quad (5)$$

All these parameters lie between 0-1 hence they can be expressed in terms of percentages. The newly introduced parameters give the better performance for higher value of LIRS and Lower value of LSRR [8-13].

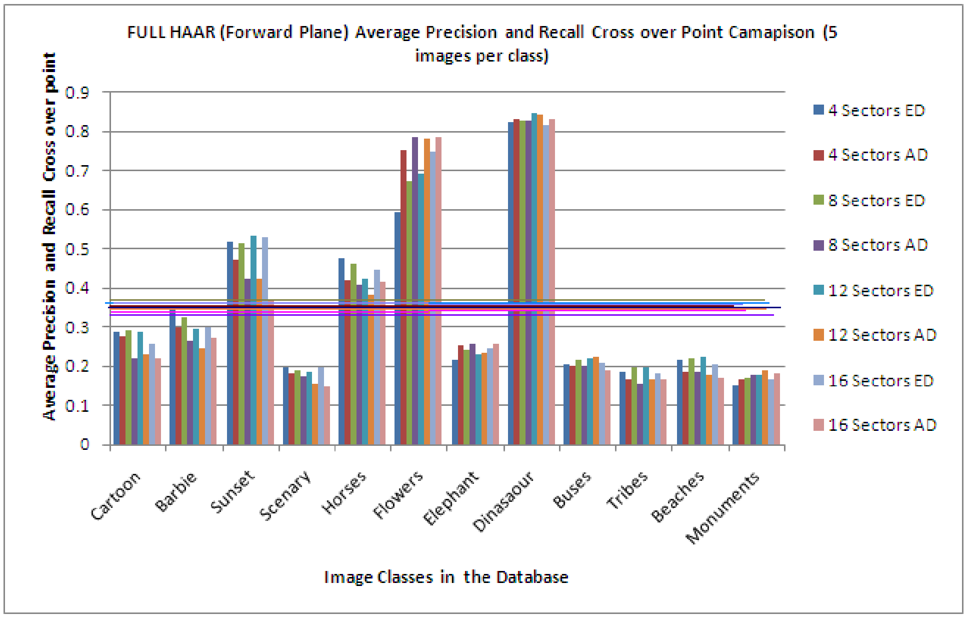


FIGURE 11: Class wise Average Precision and Recall cross over points of Forward Plane (Even) sectorization of Full Haar Wavelet with sum of Absolute Difference (AD) and Euclidean Distance (ED) as similarity measure.

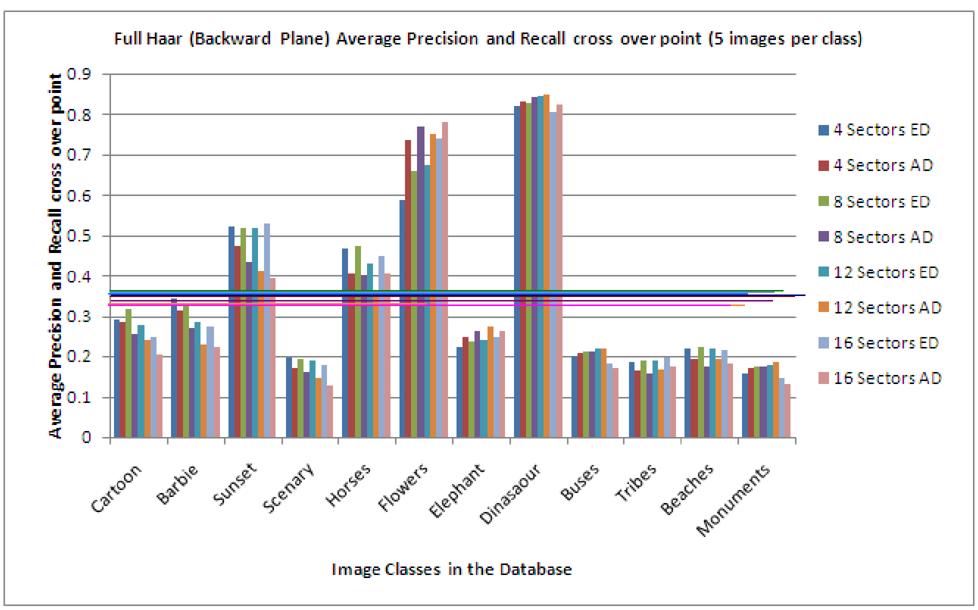


FIGURE 12: Class wise Average Precision and Recall cross over points of Backward Plane (Odd) sectorization of Full Haar Wavelet with Absolute Difference (AD) and Euclidean Distance (ED) as similarity measure.

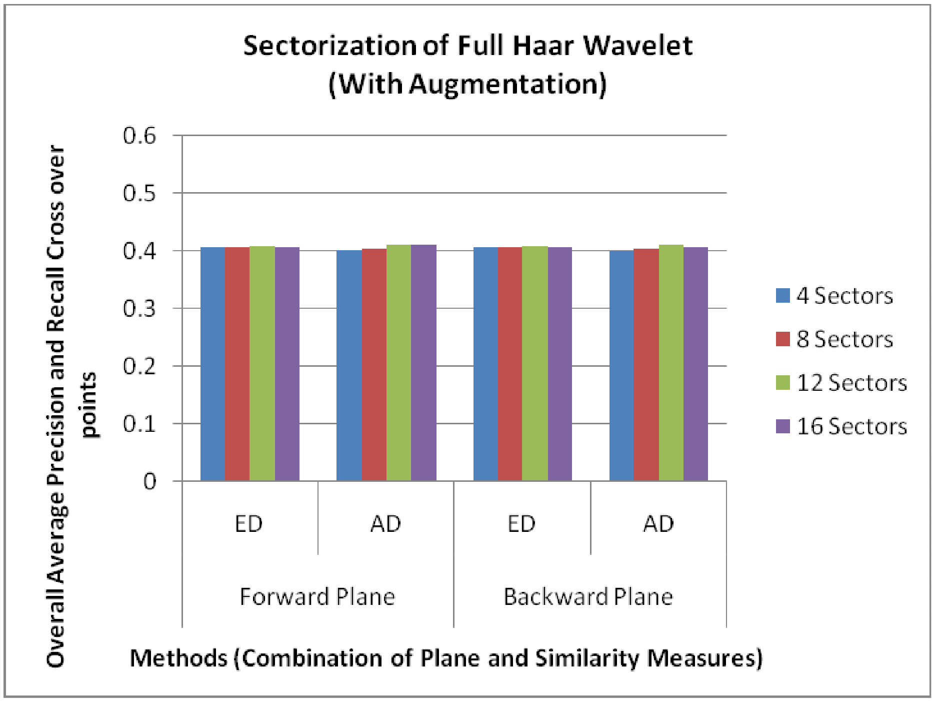


FIGURE 13: Comparison of Overall Precision and Recall cross over points of sectorization of Full Haar Wavelet with sum of Absolute Difference (AD) and Euclidean Distance (ED) as similarity measure.

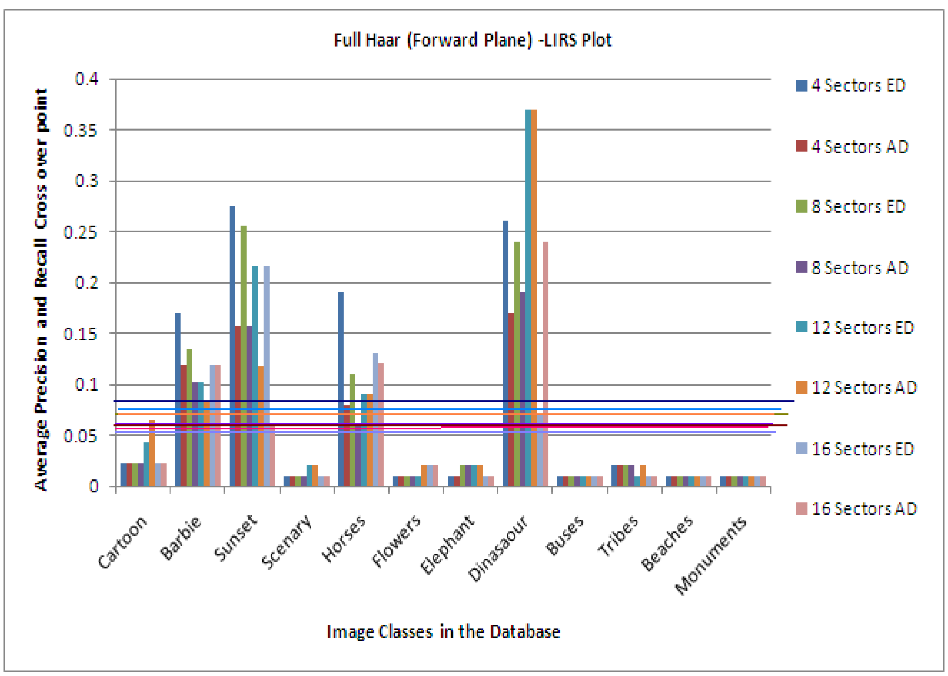


FIGURE 14: The LIRS Plot of sectorization of forward plane of Full Haar transformed images . Overall Average LIRS performances (Shown with Horizontal lines :0.082 (4 Sectors ED), 0.052 (4 Sectors AD), 0.071(8 Sectors ED), 0.051(8 Sectors AD), 0.075(12 Sectors ED), 0.069(12 Sectors AD), 0.053(16 Sectors ED), 0.053(16 Sectors AD)).

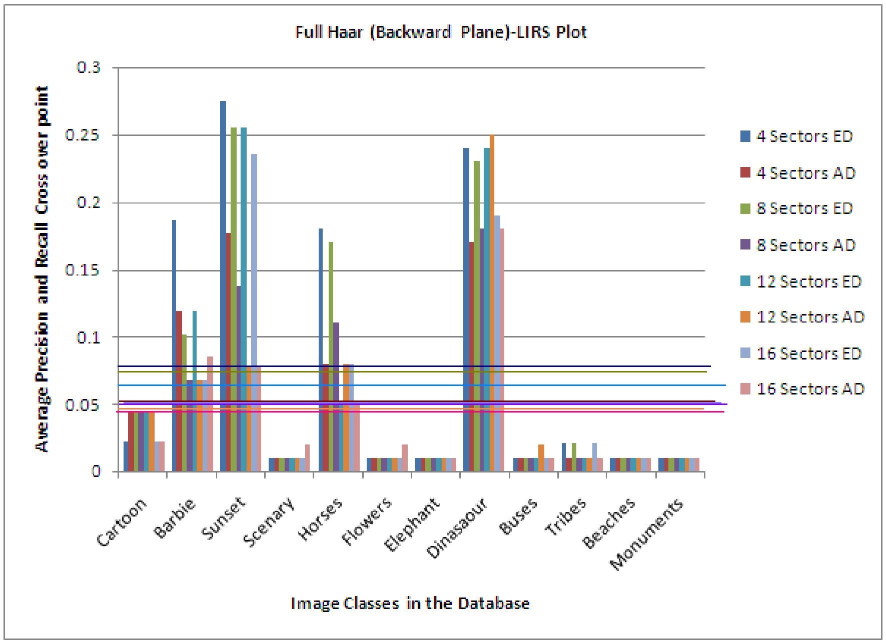


FIGURE 15: The LIRS Plot of sectorization of Backward plane of Full Haar transformed images . Overal Average LIRS performances (Shown with Horizontal lines :0.081 (4 Sectors ED), 0.054 (4 Sectors AD), 0.073(8 Sectors ED), 0.050(8 Sectors AD), 0.064(12 Sectors ED), 0.049(12 Sectors AD), 0.056(16 Sectors ED), 0.042(16 Sectors AD)).

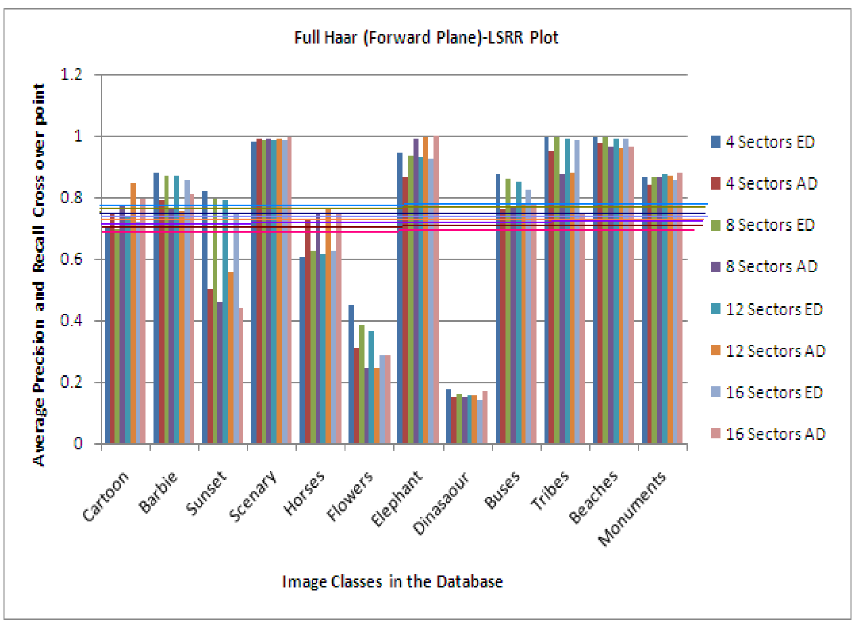


FIGURE 16: The LSRR Plot of sectorization of forward plane of Full Haar transformed images . Overall Average LSRR performances (Shown with Horizontal lines :0.77 (4 Sectors ED), 0.71 (4 Sectors AD), 0.76(8 Sectors ED), 0.71(8 Sectors AD), 0.76(12 Sectors ED), 0.73(12 Sectors AD), 0.74(16 Sectors ED), 0.71(16 Sectors AD)).

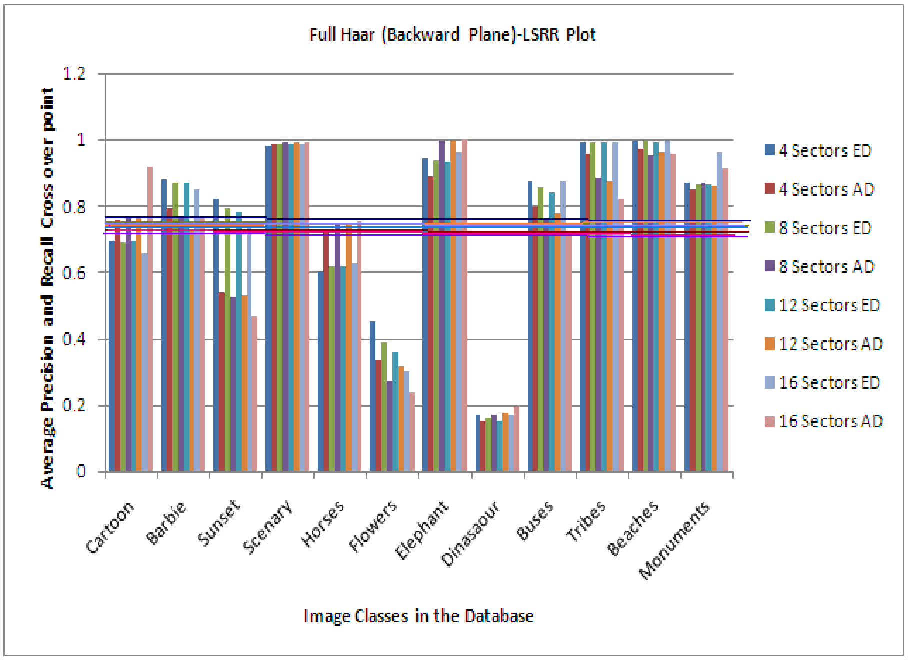


FIGURE 17: The LSRR Plot of sectorization of backward plane of Full Haar transformed images . Overall Average LSRR performances (Shown with Horizontal lines :0.77(4 Sectors ED), 0.729 (4 Sectors AD), 0.76(8 Sectors ED), 0.725(8 Sectors AD), 0.756(12 Sectors ED), 0.726(12 Sectors AD), 0.759(16 Sectors ED), 0.727(16 Sectors AD)).

5. CONCLUSION

The work experimented on 1055 image database of 12 different classes discusses the performance of sectorization of Full Haar wavelet transformed color images for image retrieval. The work has been performed with both approaches of sectorization of forward (even) plane and backward (odd) planes. The performance of the methods proposed checked with respect to various sector sizes and similarity measuring approaches namely Euclidian distance and sum of absolute difference. We calculated the average precision and recall cross over point of 5 randomly chosen images of each class and the overall average is the average of these averages. The observation is that sectorization of both planes of full Haar wavelet transformed images give 40% of the overall average retrieval of relevant images as shown in the Figure 13. The class wise plot of these average precision and recall cross over points as shown in Figure 11 and Figure 12 for both approaches depicts that the retrieval performance varies from class to class and from method to method. Few classes like sunset, horses flower and dinosaur has performance above the average of all methods as shown by horizontal lines. New parameter LIRS and LSRR gives good platform for performance evaluation to judge how early all relevant images is being retrieved (LSRR) and it also provides judgement of how many relevant images are being retrieved as part of first set of relevant retrieval (LIRS).The value of LIRS must be minimum and LSRR must be minimum for the particular class if the overall precision and recall cross over point of that class is maximum. This can be clearly seen in Figures 14 to Figure 17. This observation is very clearly visible for dinosaur class however the difference of LIRS and LSRR of other classes varies. The sum of absolute difference as similarity measure is recommended due to its lesser complexity and better retrieval rate performance compared to Euclidian distance.

6. REFERENCES

1. Kato, T., “Database architecture for content based image retrieval in Image Storage and Retrieval Systems” (Jambardino A and Niblack W eds),Proc SPIE 2185, pp 112-123, 1992.

2. Ritendra Datta,Dhiraj Joshi,Jia Li and James Z. Wang, " Image retrieval:Idea,influences and trends of the new age",ACM Computing survey,Vol 40,No.2,Article 5,April 2008.

3. Ch.srinivasa rao,S. srinivas kumar,B.N.Chaterjii, "content based image retrieval using contourlet transform", ICGST-GVIP Journal, Vol.7 No. 3, Nov2007.

4. John Berry and David A. Stoney "The history and development of fingerprinting," in Advances in Fingerprint Technology, Henry C. Lee and R. E. Gaensslen, Eds., pp. 1-40. CRC Press Florida, 2nd edition, 2001.

5. H.B.Kekre, Archana Athawale and Dipali sadavarti, "Algorithm to generate Kekre's Wavelet transform from Kekre's Transform", International Journal of Engineering, Science and Technology, Vol.2No.5,2010 pp.756-767.

6. H. B. Kekre, Dhirendra Mishra, "Digital Image Search & Retrieval using FFT Sectors" published in proceedings of National/Asia pacific conference on Information communication and technology(NCICT 10) 5TH & 6TH March 2010.SVKM'S NMIMS MUMBAI

7. H.B.Kekre, Dhirendra Mishra, "Content Based Image Retrieval using Weighted Hamming Distance Image hash Value" published in the proceedings of international conference on contours of computing technology pp. 305-309 (Thinkquest2010) 13th & 14th March 2010.

8. H.B.Kekre, Dhirendra Mishra, "Digital Image Search & Retrieval using FFT Sectors of Color Images" published in International Journal of Computer Science and Engineering (IJCSE) Vol. 02,No.02,2010,pp.368-372 ISSN 0975-3397 available online at http://www.enggjournals.com/ijcse/doc/IJCSE10-02-02-46.pdf

9. H.B.Kekre, Dhirendra Mishra, "CBIR using upper six FFT Sectors of Color Images for feature vector generation" published in International Journal of Engineering and Technology(IJET) Vol. 02, No. 02, 2010, 49-54 ISSN 0975-4024 available online at http://www.enggjournals.com/ijet/doc/IJET10-02-02-06.pdf

10. H.B.Kekre, Dhirendra Mishra, "Four Walsh transform sectors feature vectors for image retrieval from image databases", published in international journal of computer science and information technologies (IJCSIT) Vol. 1 (2) 2010, 33-37 ISSN 0975-9646 available online at http://www.ijcsit.com/docs/vol1issue2/ijcsit2010010201.pdf

11. H.B.Kekre, Dhirendra Mishra, "Performance comparison of four, eight and twelve Walsh transform sectors feature vectors for image retrieval from image databases", published in international journal of Engineering, science and technology(IJEST) Vol.2(5) 2010, 1370-1374 ISSN 0975-5462 available online at http://www.ijest.info/docs/IJEST10-02-05-62.pdf

12. H.B.Kekre, Dhirendra Mishra, " density distribution in Walsh Transform sectors ass feature vectors for image retrieval", published in international journal of compute applications (IJCA) Vol.4(6) 2010, 30-36 ISSN 0975-8887 available online at http://www.ijcaonline.org/archives/volume4/number6/829-1072

13. H.B.Kekre, Dhirendra Mishra, "Performance comparison of density distribution and sector mean in Walsh transform sectors as feature vectors for image retrieval", published in international journal of Image Processing (IJIP) Vol.4(3) 2010, ISSN http://www.cscjournals.org/csc/manuscript/Journals/IJIP/Volume4/Issue3/IJIP-193.pdf

14. H.B.Kekre, Dhirendra Mishra, "Density distribution and sector mean with zero-sal and highest-cal components in Walsh transform sectors as feature vectors for image retrieval", published in international journal of Computer science and information security (IJCSIS)

Vol.8(4) 2010, ISSN 1947-5500 available online http://sites.google.com/site/ijcsis/vol-8-no-4-jul-2010

15. Arun Ross, Anil Jain, James Reisman, "A hybrid fingerprint matcher," Int'l conference on Pattern Recognition (ICPR), Aug 2002.

16. H.B.Kekre, Dhirendra Mishra, "DCT sectorization for feature vector generation in CBIR", International journal of computer application (IJCA),Vol.9, No.1,Nov.2010,ISSN:1947-5500 http://ijcaonline.org/archives/volume9/number1/1350-1820

17. A. M. Bazen, G. T. B.Verwaaijen, S. H. Gerez, L. P. J. Veelenturf, and B. J. van der Zwaag, "A correlation-based fingerprint verification system," Proceedings of the ProRISC2000 Workshop on Circuits, Systems and Signal Processing, Veldhoven, Netherlands, Nov 2000.

18. H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, "Image Retrieval using Color-Texture Features from DST on VQ Codevectors obtained by Kekre's Fast Codebook Generation", ICGST International Journal on Graphics, Vision and Image Processing (GVIP), Available online at http://www.icgst.com/gvip

19. H.B.Kekre, Sudeep D. Thepade, "Using YUV Color Space to Hoist the Performance of Block Truncation Coding for Image Retrieval", IEEE International Advanced Computing Conference 2009 (IACC'09), Thapar University, Patiala, INDIA, 6-7 March 2009.

20. H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Augmented Block Truncation Coding Techniques", ACM International Conference on Advances in Computing, Communication and Control (ICAC3-2009), pp.: 384-390, 23-24 Jan 2009, Fr. Conceicao Rodrigous College of Engg., Mumbai. Available online at ACM portal.

21. H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, "DST Applied to Column mean and Row Mean Vectors of Image for Fingerprint Identification", International Conference on Computer Networks and Security, ICCNS-2008, 27-28 Sept 2008, Vishwakarma Institute of Technology, Pune.

22. H.B.Kekre, Vinayak Bharadi, "Walsh Coefficients of the Horizontal & Vertical Pixel Distribution of Signature Template", In Proc. of Int. Conference ICIP-07, Bangalore University, Bangalore. 10-12 Aug 2007.

23. J. L. Walsh, "A closed set of orthogonal functions" American Journal of Mathematics, Vol. 45, pp.5-24,year 1923.

24. H.B.Kekre, Dhirendra Mishra, "DST Sectorization for feature vector generation", Universal journal of computer science and engineering technology (Unicse),Vol.1, No.1Oct 2010 available at http://www.unicse.oeg/index.php?option=comcontent and view=article&id=54&itemid=27

A New Wavelet Based Digital Watermarking Method for Authenticated Mobile Signals

Dr. B. Eswara Reddy eswarcsejntu@gmail.com
Associate Professor & HOD, Dept. of CSE,
JNTUA College of Engineering
Anantapur, A.P., India.

P. Harini hariniphd@yahoo.com
Professor & HOD, Department. of IT,
St. Anns College of Engineering and Technology
Chirala, A.P., India.

S. MaruthuPerumal maruthumail@gmail.com
Research Scholar, Dr. M.G.R. University Chennai.
Associate Professor, IT Godavari Institute of
Engineering and Technology, Rajahmundry,
A.P., India.

Dr. V. VijayaKumar vakulabharanam@hotmail.com
Professor & Dean, Department of Computer Sciences
Head - Srinivasa Ramanujan Research Forum-SRRF,
Godavari Institute of Engineering and Technology, Rajahmundry,
A.P., India.

Abstract

The mobile network security is becoming more important as the number of data being exchanged on the internet increases. The growing possibilities of modern mobile computing environment are potentially more vulnerable to attacks. As a result, confidentiality and data integrity becomes one of the most important problems facing Mobile IP (MIP). To address these issues, the present paper proposes a new Wavelet based watermarking scheme that hides the mobile signals and messages in the transmission. The proposed method uses the successive even and odd values of a neighborhood to insert the authenticated signals or digital watermark (DW). That is the digital watermark information is not inserted in the adjacent column and row position of a neighborhood. The proposed method resolves the ambiguity between successive even odd gray values using LSB method. This makes the present method as more simple but difficult to break, which is an essential parameter for any mobile signals and messages. To test the efficacy of the proposed DW method, various statistical measures are evaluated, which indicates high robustness, imperceptibility, un-ambiguity, confidentiality and integrity of the present method.

Keywords: Mobile Network, Wavelet Transform, Even-Odd Method, Statistical Measures.

1. INTRODUCTION

The rapid growth of the Internet increased the access to multimedia data tremendously [1]. The development of digital multimedia is demanding as an urgent need for protect multimedia data in internet. Digital watermarking technique provides copyright protection for digital data [23,2,3]. The digital watermarking technique is proposed as a method to embed perceptible or imperceptible signal into multimedia data for claiming the ownership. A digital watermark is a piece of information which is embedded in the digital media and hidden in the digital content in such a way that it is inseparable from its data. This piece of information known as watermark, a tag, or label into multimedia object such that the watermark can be detected or extracted later to make an assertion about the object. The object may be an image, audio, video, or text [4].

Each watermarking application has its own requirements that determine the required attributes of the watermarking system and drive the choice of techniques used for embedding the watermark [5]. This demand has been lately addressed by the emergence of a variety of watermarking methods. Such methods target towards hiding an imperceptible and undetectable signal in the original data, which conveys copyright information about the owner or authorized user. Data hiding usually involves the use of secret keys possessed only by owners or authorized users. In order to verify the multimedia content owner and thus protect his copyrights, detection is performed to test whether the material in question is watermarked with his own secret key [19, 20]. Recent research trend in watermarking technique is focusing more on image data [6, 7, 8, 24, 25]. But watermarking is not limited to only images; but there are also watermarking techniques for audio [9, 10], video [26, 27, 11], and text [12, 13, 14, 15, 16, 17] data. Watermarking for black and white text data; e.g., electronic documents and manuscripts, is so-called binary watermarks [18]. Watermarks and Watermarking techniques can be divided into various categories. The watermarks can be applied either in spatial domain or frequency domain. The spatial domain watermarking schemes have less computational overhead compared with frequency domain schemes.

Authentication is an important security requirement, which conventionally requires the identity of a user or some other identifying information in terms of signals of a node [22]. On the other hand, users and consumers are becoming increasingly concerned about their privacy, and the risks (such as identity theft) of leaving any form of digital trail, when making electronic transactions. Hence, given a choice, users may well prefer to interact with service providers anonymously (or pseudonymously). Under these circumstances, it may in fact be undesirable to authenticate the identity of a user. Hence, to preserve the privacy of a user, or the routing signals or any other signals of a mobile IP are needed to be authenticated. Achieving security and privacy concurrently, whilst protecting the interests of both users and service providers remains an open problem. So, privacy preserving authentication schemes may be needed to be devised. Many Content Distribution Protection (CDP) schemes (e.g. Buyer– Seller watermarking and asymmetric fingerprinting) have since been proposed to address the problem of illegal distribution of copyrighted content. All of the existing CDP schemes heavily rely on (online) Trusted Third Party in one way or another to achieve the desired security objectives. This requirement for (online) Trusted Third Party in existing Buyer Seller Watermarking (BSW) and Asymmetric Fingerprinting (AF) schemes, either, to generate the buyer watermarks, or to provide pseudonyms for buyers, represents a major constraint. To provide a better security for a mobile environment, the present paper has carried out an in depth study on the existing Digital Watermarking algorithms and outlined a new spatial domain watermarking method. The present paper is organized as follows. Section 2 deals with the introduction of wavelet transform, section 3 introduces the proposed new mechanism, section 4 deals with results and discussions and section 5 deals with conclusions.

2. INTRODUCTION TO WAVELET

The wavelet transformation is a mathematical tool for decomposition. The wavelet transform is identical to a hierarchical sub band system [18, 19], where the sub bands are logarithmically spaced in frequency. The basic idea of the DWT for a two-dimensional image is described as follows. An image is first decomposed into four parts based on frequency sub bands, by critically sub sampling horizontal and vertical channels using sub band filters and named as Low-Low (LL), Low-High (LH), High-Low (HL), and High- High (HH) sub bands as shown in Figure 1. To obtain the next coarser scaled wavelet coefficients, the sub band LL is further decomposed and critically sub sampled. This process is repeated several times, which is determined by the application at hand. The block diagram of this process is shown in Figure 1. Each level has various bands information such as low–low, low–high, high–low, and high–high frequency bands. Furthermore, from these DWT coefficients, the original image can be reconstructed. This reconstruction process is called the inverse DWT (IDWT). If C[m,n] represents an image, the DWT and IDWT for

C[m,n] can similarly be defined by implementing the DWT and IDWT on each dimension and separately.

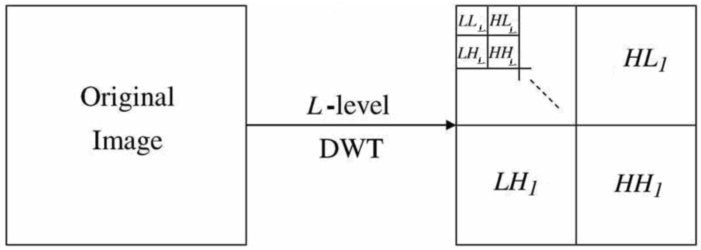


FIGURE 1: Representation of n-Levels of Wavelet Transform

The Haar wavelet [21] is the first known wavelet and was proposed in 1909 by Alfred Haar. Haar used these functions to give an example of a countable orthonormal system for the space of square-integrable functions on the real line. The Haar wavelet's scaling function coefficients are h{k}={0.5, 0.5}and wavelet function coefficients are g{k}={0.5, -0.5}.In the wavelet transform [28],[29],[30],[31],[32],[33],[34],[35],[36],[37] an image signal can be analyzed by passing it through an analysis filter bank followed by a decimation operation. This analysis filter bank consists of a low pass and a high pass filter at each decomposition stage. When the signal passes through these filters it splits into two bands. The low pass filter, which corresponds to an averaging operation, extracts the coarse information of a signal. The high pass filter, which corresponds to a differencing operation, extracts the detail information of the signal. The output of the filtering operations is then decimated by two [28], [34], [36].Because of decimation the total size of the transformed image is same as the original image, which is shown in the figure 2 as horizontal wavelet transform. Then, it is followed by filtering the sub image along the y-dimension and decimated by two. Finally, the image splits into four bands denoted by low-low (LL1), high-low (HL1),low-high (LH1) and high-high (HH1) after one-level decomposition as depicted in figure 3. The sub bands labeled, LL1 HL1, LH1, and HH1 represent the finest scale wavelet coefficients. To obtain the next coarser scaled wavelet coefficients, the sub band LL1 is further decomposed and critically sub sampled. This process is repeated several times, which is determined by the application at hand. Furthermore, from these DWT coefficients, the original image can be reconstructed. This reconstruction process is called the inverse DWT (IDWT). If I[m,n]represents an image, the DWT and IDWT for I[m,n] can be similarly defined by implementing the DWT and IDWT on each dimension and separately. Figure 4 shows second level of filtering. This process of filtering the image is called 'Pyramidal decomposition' of image. Figure 5 shows the results of different levels of Haar wavelet decomposition of Lena image.

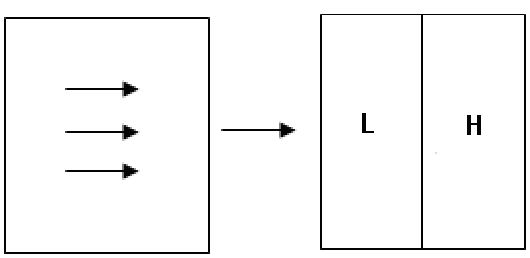


FIGURE 2: Horizontal Wavelet Transform.

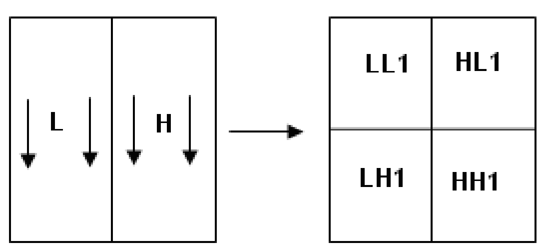


FIGURE 3: Vertical Wavelet Transform

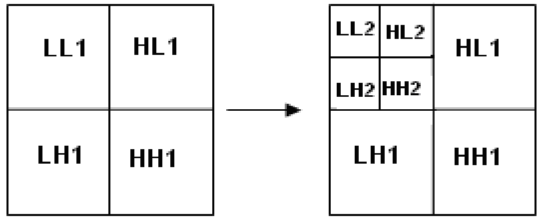


FIGURE 4: Second Level Wavelet Transform.

(a) (b) (c) (d)

FIGURE 5: (a) Original Lena Image (b) Level-1 Haar Wavelet Transformed Lena Image (c) Level-2 Haar Wavelet Transformed Lena Image (d) Level-3 Haar Wavelet Transformed Lena Image.

3. METHODOLOGY

The proposed method embeds the given mobile signals (watermark) in the n level decomposed LL subband of the original image. The proposed algorithm is given in the following steps.

Step 1: Apply the n- level Haar wavelet transform on the input image and obtain the n-level LL subband image.

Step 2: Let the n-level LL subband image is represented by C(x,y). The n-level LL subband image is divided into non overlapped windows of size 5×5. Let the window be Wi i=1, 2 …….. n. The Wi will be having a central pixel and 24 neighboring pixels as shown in Figure 6. All windows are assumed logically to start with window coordinates of (0, 0). (i,j) represents the coordinate position and CP represents Central Pixel.

Wi=

(i,j)	(i,j+1)	(i,j+2)	(i,j+3)	(i,j+4)
(i+1,j)				
(i+2,j)		CP(i+2,j+2)		
(i+3,j)				
(i+4,j)				(i+4,j+4)

FIGURE 6: A 5×5 window with coordinate position

Step 3: Arrange the gray level values in the ascending order form along with their coordinate value, Pi (xi, yj), Pi+1 (xi+l, yj+l) …….; here Pi(xi,yi) denotes the gray level value of the location (xi, yj).

Step 4: Consider the successive even (ei) and odd gray values (ei+1) as same after sorting. Where ((ei+1)-ei) is always one and ei<ei+1. If two or more pixels of the window have the same gray level values or if they are successive even and odd values of the window then the least coordinated value of row and column will be treated as least value. The watermark bit will be embedded in the ascending order of gray level values of the considered window 5x5 on the least x-coordinate and y-coordinate position. This process is explained with an example.

Step 5: Convert each character of the watermark in to a 12 bit character by appending the MOD 9 value of each character.

Step 6: Insert the bits of watermark in to the identified pairs in ascending order of step-4.

Step 7: Repeat the process for each 5×5 non-overlapped n- level wavelet based window.

Step 8: Apply n-level inverse wavelet transformation to obtain the watermarked image.

The detailed explanation of digital watermarking embedding process is given below with an example. Any image is represented by a two dimensional array of values $f(x_i,y_j)$ where $0 \leq (i, j) \leq N$. The present paper divides the image into non overlapped window of a predefined size. Any window of size m x m will be having m^2 pixels. Figure 7 shows the gray level values of an 5x5 window. The Table1 gives the sorted list of gray level values with the co-ordinate position of figure7 as specified in step 3. The present method considers the pair of values (80,81), (78,79),(76,77), and (74,75) as same because they are successive even and odd gray values(ei, (ei+1)) as specified in step 4. By this method the successive even and odd values of the figure 7 are treated as same and the watermark bit is inserted based on the least x-coordinate and y-coordinate position. The order of embedding bits of the figure 7 is shown in table 2.

78	75	79	80	81
81	80	81	81	80
76	75	76	75	80
78	77	79	78	75
80	81	75	74	73

FIGURE 7: Gray level Values of the image 5x5

Coordinate Positions		Pi(xi,yi)
xi	yi	
4	4	73
4	3	74
0	1	75
2	1	75
2	3	75
3	4	75
4	2	75
2	0	76
2	2	76
3	1	77
0	0	78
3	0	78
3	3	78
0	2	79
3	2	79
0	3	80
1	1	80
1	4	80
2	4	80
4	0	80
0	4	81
1	0	81
1	2	81
1	3	81
4	1	81

TABLE 1: Sorted gray level values with the coordinate position of figure 7

Coordinate Positions		Pi(xi,yi)
xi	yi	
4	4	73
0	1	75
2	1	75
2	3	75
3	4	75
4	2	75
4	3	74
2	0	76
2	2	76
3	1	77
0	0	78
0	2	79
3	0	78
3	2	79
3	3	78
0	3	80
0	4	81
1	0	81
1	1	80
1	2	81
1	3	81
1	4	80
2	4	80
4	0	80
4	1	81

TABLE 2: Embedding order of watermark bits as specified in step 4

In the proposed method, successive even and odd values are treated as same but not successive odd and even values. Because an even number will have always a zero in the LSB, even by embedding a '1' in the LSB, it's value will be incremented by one at most. In the same way an odd number will have always a one in the LSB, even by embedding a '0' in the LSB its value will be at most decremented by one. That is the odd values will never increment by 1 after embedding the digital watermark bit. And the even values will never decrement by 1 after embedding the digital watermark bit. Therefore always the maximum difference between successive even and odd values will be one after embedding the digital watermark bit. Where as the maximum difference between successive odd and even values will be two after inserting the digital watermark bit. For this reason, the successive even and odd values of a neighborhood are treated as same in the proposed approach. This property removes the ambiguity in the extraction process of the watermark bits, based on ascending order of the window.

The method has used various quality measures for the watermarked image, like Mean Square Error (MSE), Root Mean Square Error (RMSE), Peak Signal to Noise Ratio (PSNR), Signal to Noise Ratio (SNR), Root Signal to Noise Ratio (RSNR) and Normalized Correlation Coefficient

(NCC) given in Equations (1.1) to (1.6). The embedding distortion performance is usually measured by the Mean Square Error and Root Mean Square Error as given by the equation (1.1) and (1.2). MSE and RMSE for the image should be as low as possible. A better perceptual quality of the image is possible for lower values of MSE and RMSE. Signal to Noise Ratio (SNR) and Root SNR (RSNR) measures, estimates the quality of the reconstructed image compared with an original image. The fundamental idea is to compute the value which reflects the quality of the reconstructed image. Reconstructed image with higher metric are judged as having better quality. SNR and RSNR can be defined as in equation (1.3) and (1.4). A higher value of SNR and RSNR indicates the better quality of the reconstructed image. To study the quality of watermarked image, the peak signal to noise ratio PSNR is used. In general, a processed image is acceptable to the human eyes if its PSNR is greater than 30 dB. The larger the PSNR, the better is the image quality. The PSNR is defined as given by the equation (1.5). To verify the robustness of any digital watermarking method, Normalized Cross Correlation (NCC) is used. NCC is an important performance parameter in any extracting module. NCC is defined in the equation (1.6). Using NCC, the similarity value about 0.75 or above is considered as acceptable.

$$MSE = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (f(i,j) - g(i,j))^2 \tag{1.1}$$

$$RMSE = \sqrt{\frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (f(i,j) - g(i,j))^2} \tag{1.2}$$

$$SNR = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} g(i,j)^2}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (g(i,j) - f(i,j))^2} \tag{1.3}$$

$$RSNR = \sqrt{\frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} g(i,j)^2}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (g(i,j) - f(i,j))^2}} \tag{1.4}$$

$$PSNR = 10 * \log_{10}\left(\frac{255^2}{MSE}\right) \tag{1.5}$$

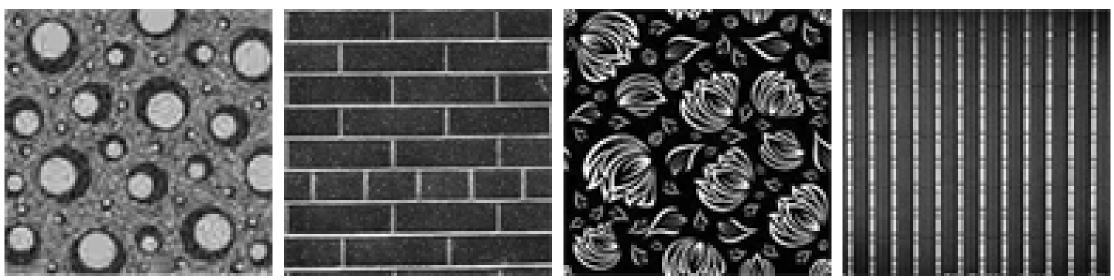
where f(i,j) and g(i,j) are the original image and watermarked image with coordinate position (i,j).

$$NCC = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} g(i,j) \times f(i,j)}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} g(i,j)^2} \tag{1.6}$$

where f(i,j) and g(i,j) are the original watermark and extracted watermark.

4. RESULTS AND DISCUSSION

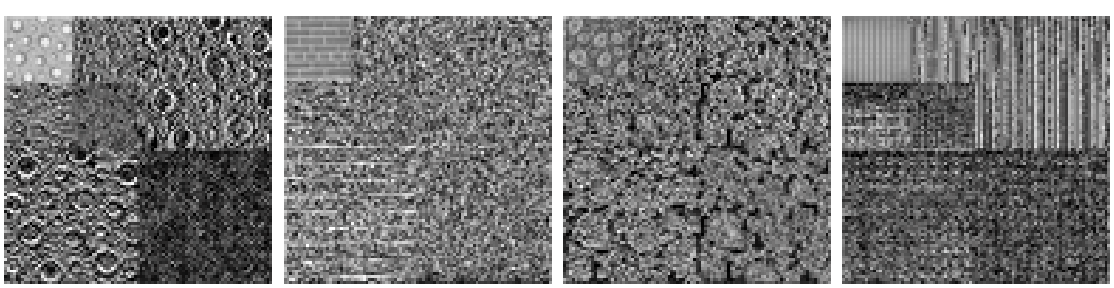
For the experimental analysis on the proposed method different images of size 256×256 are taken. The cover images considered in the proposed method are Circle, Brick, Curve, Line, Lena and Cameraman images which are shown from Figure 7(a) to 7(f) respectively. The proposed method is applied on 2-level Haar wavelet decomposed images. In the wavelet based watermarked image, a set of 16 characters “srinivasaramanuj” is embedded. The set of 16 characters is chosen because the present study founds that no mobile signal is more than 16 characters. The Figure 8(a) to 8(f) shows the 2- level proposed method based watermark (16 bit character) embedded images. The Figure 9(a) to 9(f) shows the reconstructed watermarked image. The values of above discussed quality parameters are calculated on the proposed watermarked images and are listed in the Table 3. The Table 3 clearly indicates the efficacy of the proposed scheme since the quality measures of watermarked images falls with in the good range i.e. the proposed method produces minimum values for MSE, RMSE, NCC and higher values for SNR, PSNR, RSNR.



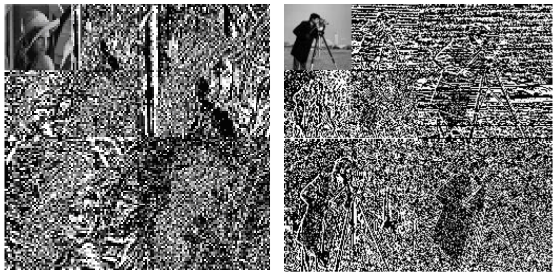
(a) (b) (c) (d)

(e) (f)

FIGURE 7: Original Images (a) Circle Image (b) Brick Image (c) Curve Image (d) Line Image (e) Lena Image (f)Cameraman Image

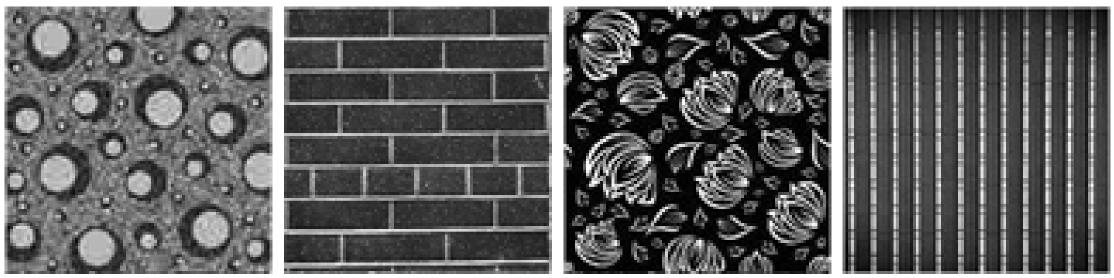


(a) (b) (c) (d)



(e) (f)

FIGURE 8: Two level wavelet transformed images (a) Circle Image (b) Brick Image (c) Curve Image (d) Line Image (e) Lena Image (f) Cameraman Image



(a) (b) (c) (d)

(e) (f)

FIGURE 9: Reconstructed watermarked Images with the proposed method (a) Circle Image (b) Brick Image (c) Curve Image (d) Line Image (e) Lena Image (f) Cameraman Image

S.No	Image Name	MSE	RMSE	SNR	RSNR	PSNR	NCC
1	Circle	0.0237	0.1539	212880.0000	461.3800	64.3870	0.9997
2	Brick	0.0293	0.1712	265050.0000	514.8300	63.4630	0.9998
3	Curve	0.0242	0.1555	378250.0000	615.0200	64.2980	0.9999
4	Line	0.0222	0.1491	400070.0000	632.5100	64.6640	0.9999
4	Lena	0.0247	0.1629	396050.0000	589.2600	64.7623	0.9997
5	Cameraman	0.0241	0.1633	387904.0000	599.3200	63.3452	0.9996

TABLE 3: Quality Measures for the watermarked images

5. CONCLUSIONS

The proposed watermarking method is applied on 2- level decomposed images to hide the mobile signals or messages for high authentication and security. The experimental result with various statistical parameters indicates high robustness, imperceptibility, un-ambiguity, confidentiality and integrity of the proposed digital watermarking method which are the essential features for mobile transmissions. The novelty of the proposed method is it embeds the information in a non linear order based on the values and position of a window. Moreover each 8-bit character is represented as a 12 bit character in the proposed method. This makes the proposed method as more secured when compared to the other methods especially for transmission of mobile signals. To reduce the time limit, the present method can directly apply on a original image in which case the robustness, imperceptibility, confidentiality and integrity will be degraded little bit.

Acknowledgements
The authors would like to express their cordial thanks to Dr. K.V.V. Satya Narayana Raju, Chairman for providing facility to work at the advanced labs of SRRF-GIET and Sri. K. Sasi Kiran Varma, Managing Director, Chaitanya group of Institutions for providing moral support and encouragement towards research. Authors would like to thank Dr. G.V.S. Anantha Lakshmi and M. Radhika Mani for their invaluable suggestions and constant encouragement that led to improvise the presentation quality of the paper.

REFERENCES

[1] Cox, J. Kilian, F. Leighton, and T. Shamoon, “Secure spread spectrum watermarking for multimedia,” IEEE Trans. Image Processing, vol. 6, pp. 1673–1687, Dec. 1997.

[2] Houng.Jyh Wang and c.c, Jay Kuo, “Image protection via watermarking on perceptually significant wavelet coefficient”, IEEE 1998 workshop on multimedia signal processing, Redondo Beach, CA, Dec, 7-9,1998.

[3] S.Craver, N. Memon, B.L and M.M Yeung “Resolving rightful ownerships with invisible watermarking techniques: limitations, attacks, and implication”, IEEE Journal on selected Areas in communications. Vol.16 Issue:4, May 1998, pp 573-586.

[4] S.Katzenbeisser and F.A.P. Petitcolas, Information Hiding Techniques for steganography and Digital Watermarking. Norwood,MA: Artech House,2000.

[5] Ravik. Sharma, and SteveDecker,” Practical Challenges for Digital Watermarking Applications”, Proc. IEEE,2001.

[6] G. Braudaway, “Results of attacks on a claimed robust digital image watermark,” In Proceedings of the IS&T/SPIE Symposium onOptical Security and Counterfeit Deterrence Techniques, SPIE, pp. 122-131, 1998.

[7] G. Braudaway, K. Magerlein, and F. Mintzer, “Color Correct Digital Watermarking of Images,” U.S. Patent 5 530 759, June 25, 1996.

[8] H. Gladney, F. Mintzer, and F. Schiattarella, “Safeguarding digital library contents and users: Digital images of treasured antiquities,” DLIB Mag., 1997;

[9] L. Boney, A.H. Tewfik, and K.N. Hamdy, “Digital Watermarks for Audio Signals,” Proceedings of MULITMEDIA’96, pp. 473- 480, 1996.

[10] M. Swanson, B. Zhu, A.H. Tewfik, and L. Boney, “Robust Audio Watermarking using Perceptual Masking,” Signal Processing, (66), pp. 337-355, 1998.

[11] M. Swanson, B. Zhu, A.H. Tewfik, “Object-based Transparent Video Watermarking,” Proceedings 1997 IEEE Multimedia Signal Processing Workshop, pp. 369-374, 1997.

[12] J. Brassil, S. Low, N. Maxemchuk, and L. O'Gorman, "Electronic Marking and Identification Techniques to Discourage Document Copying," IEEE J. Select. Areas Commun., Vol. 13, pp.1495-1504, Oct. 1995.

[13] S. Low, N. Maxemchuk, J. Brassil, and L. O'Gorman, "Document Marking and Identification Using Both Line and Word Shifting," In Proc. Infocom'95, 1995. (Available WWW: http://www.research.att/com:80/lateinfo/projects/ecom.html.)

[14] N. Maxemchuk, "Electronic Document Distribution," AT&T Tech. J., pp. 73-80, Sept. 1994.

[15] SysCoP, http://www.mediasec.com/products/products.html.

[16] P. Vogel, "System For Altering Elements of a Text File to Mark Documents," U.S. Patent 5 338 194, Feb. 7, 1995.

[17] J. Zhao and Fraunhofer Inst. For Computer Graphics (1996). [Online]. Available WWW: http://www.igd.fhg.de/~zhao/zhao.html.

[18] A. Shamir, "Method and Apparatus for Protecting Visual Information With Printed Cryptographic Watermarks," U.S. Patent 5488664, Jan. 1996.

[19] Antonini, M., Barlaud, M., Mathieu, P., Daubechies, I., 1992. Image coding using wavelet transform. IEEE Trans. Image Process. 1 (2), 205–220.

[20] Daubechies Ingrid., 1992. Ten Lectures on Wavelets. Rutgers University and AT&T Laboratories.

[21] Soman K.P., Ramachandran K.I., "Insight into Wavelets from theory to practice," Prentice Hall of India Pvt. Limited, 2004.

[22] Adrian Leung, Yingli Sheng, Haitham Cruickshank, "The security challenges for mobile ubiquitous services," Information security technical report, Elsevier, Vol. 12, pp.162 – 171, 2007.

[23] N. Nikolaidis and I. Pitas, "Copy right Protection of images using robust digital signatures", in proceeding , IEEE International Conferences on Acoustics, Speech and signal processing , Vol.4, May 1996,pp. 2168-2171.

[24] F. Mintzer, G. Braudaway, and M. Yeung, "Effective and ineffective image watermarks,"In Proceedings of the IEEE ICIP'97, pp. 9-12, 1997.

[25] J. Zhao, and E. Koch, "Embedding Robust Labels into Images for Copyright Protection," Proceedings of the KnowRight'95 conference, pp. 242 – 251, 1995.

[26] M. Swanson, B. Zhu, and A.H. Tewfik, "Data Hiding for Video in Video," Proceedings of the IEEE International Conference on Image Processing 1997, Vol. II, pp. 676-679, 1997.

[27] M. Swanson, B. Zhu, and A.H. Tewfik, "Multiresolution Video Watermarking using Perceptual Models and Scene Segmentation," Proceedings of the IEEE International Conference on Image Processing 1997, Vol. II, pp. 558-561, 1997.

[28] Daubechies Ingrid "Ten Lectures on Wavelets," Society for Industrial and Applied Mathematics, 1992.

[29] Jaideva C. Goswami and Andrew K. Chan “Fundamentals of Wavelets Theory, Algorithms, and Applications,” A wiley- Interscience Publications.

[30] Michael W. Frazier “An Introduction to Wavelets through Linear Algebra,” Replika Press Pvt Ltd, Delhi, India, 1999.

[31] Nievergelt Yves “Wavelets Made Easy,” R.R. Honnelley and Sons, Harrisonburg, USA, 1999.

[32] Pankaj N. Topiwala “Wavelet Image and Video Compression”, Kluwer Academic Publishers, 1998.

[33] Prasad L. and Iyengar S.S. “Wavelet analysis with applications to image processing,” CRC press LLC, Florida, USA, 1997.

[34] Rafael C. Gonzalez and Richard E. Woods “Digital Image Processing,” 2nd ed., Pearson Education (Singapore) Pte. Ltd, Indian Branch, 2003.

[35] Raghuveer M. Rao and Ajit S. Bopardikar “Wavelet Transforms-Introduction to Theory and Applications,” Pearson Education, Published by Dorling Kindersley (India) Pvt. Ltd, 2006.

[36] Soman K.P. and Ramachandran K.I. “Insight into wavelets from theory to practice,” Prentice-Hall of India Pvt. Limited, 2004.

[37] Soman K.P. and Ramachandran K.I. “Insight into wavelets from theory to practice,” Second Edition, Prentice-Hall of India Pvt. Limited, 2006.

Fabric Textile Defect Detection, By Selecting A Suitable Subset Of Wavelet Coefficients, Through Genetic Algorithm

Narges Heidari narges_hi@yahoo.com
Department Of Computer
Islamic Azad University, ilam branch
Ilam, Iran

Reza Azmi azmi@alzahra.ac.ir
Department Of Computer
Alzahra University
Tehran, Iran

Boshra Pishgoo boshra.pishgoo@yahoo.com
Department Of Computer
Alzahra University
Tehran, Iran

Abstract

This paper presents a novel approach for defect detection of fabric textile. For this purpose, First, all wavelet coefficients were extracted from an perfect fabric. But an optimal subset of These coefficients can delete main fabric of image and indicate defects of fabric textile. So we used Genetic Algorithm for finding a suitable subset. The evaluation function in GA was Shannon entropy. Finally, it was shown that we can gain better results for defect detection, by using two separable sets of wavelet coefficients for horizontal and vertical defects. This approach, not only increases accuracy of fabric defect detection, but also, decreases computation time.

Keywords: Fabric Textile Defect Detection, Genetic Algorithm, Wavelet Coefficients.

1. INTRODUCTION

In loom industry, it's very important to control quality of textile generation. Though, loom machines improvement, decreases defects in fabric textile, but this process can't perform quite perfect [1]. In the past, human resource performed defect detection process in factories but this tedious operation, depended on their attention and experiment. So we couldn't expect high accuracy from them. In other words, according to experimental results that were shown in Ref [1], human eyesight system can recognize just 50-70% of all defects in fabric textile. In the other hand, Ref [2] claimed that human's accuracy in fabric defect detection isn't higher than 80%. So, it seems that automatic evaluation of fabric defects is very necessary.

Because of improvement in image processing and pattern recognition, automatic fabric defect detection can presents an accurate, quick and reliable evaluation of textile productions. since fabric evaluation is very important in industrial applications, Different researches were done on fabric textile, surface of wood, tile, aluminum and etc [2,3,4,5,6]. [3] presented a method for fabric defect detection based on distance difference. In this method, user can adjust some parameters, according to the type of textile. [7] evaluated effects of using regularity and local directions as two features for fabric textile defect detection. [8] used Gabor filters for feature extraction and Gaussian Mixture model for detection and classification of fabric defects. [9] used structure characters of fabrics for defect detection and [10] was done this process by using adaptive local binary patterns.

Many defect detection systems that were presented up now, work on offline mode. In this mode, if occur some defects in a part of the fabric textile because of machine's problem, all parts of the fabric will be defective. But in online mode, since defect detection operation perform on textile generation time, we can stop generation process as soon as occurs a defection. The goal of this paper is to presents a method for online processing with acceptable accuracy and speed in defect detection.

Here, we used wavelet filter for fabric defect detection. This filter puts special information in each quarter of filtered image. In two dimensional wavelet transform, initial image divides to four parts. The first part is an approximation of initial image and the other parts include components with high frequencies or details of image, so vertical, horizontal and diagonal defects can be found in part 2, 3 and 4 respectively. There is a problem for using wavelet transform in fabric defect detection. The wavelet coefficients that are suitable for a special fabric, may aren't suitable for the other one. So in this paper, we used separable wavelet coefficients sets for each fabric.

After extracting all wavelet coefficients from a perfect fabric, we found a suitable subset of these coefficients using genetic algorithm. Then applied this suitable subset to other images and gain the available defects in each of them. In fact, our final goal is to design a system that can recognize defective and perfect fabrics. This paper is organized as follows: in Section 2 and 3, we briefly introduce Wavelet transform and Genetic algorithm. Section 4 describes the proposed method for fabric defect detection and Our experimental results is shown in Section 5.

2. WAVELET TRANSFORM

Wavelet transform analyses the signal by a group of orthogonal functions in the form of $\Psi_{m,n}(x)$, that computed from translation and dilation of the mother wavelet Ψ. These orthogonal functions are shown in Equation 1.

$$\Psi_{m,n}(x) = 2^{-\frac{m}{2}}\psi(2^{-m}x - n) \tag{1}$$

m and *n* are real numbers in this equation. Analysis and synthesis formulas are shown in Equation 2 and 3 respectively.

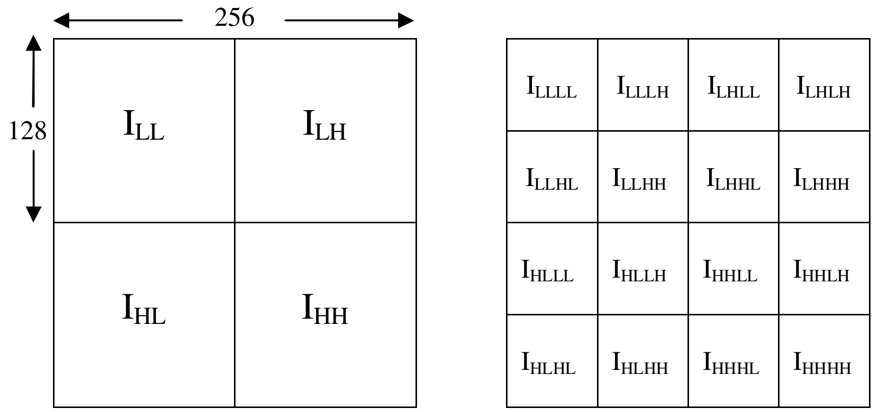
$$C_{m,n} = \int_{-\infty}^{+\infty} f(x)\,\psi_{m,n}(x)\,dx \tag{2}$$

$$f(x) = \sum_{m,n} C_{m,n}\,\psi_{m,n(x)} \tag{3}$$

Coefficients of this transform are obtained recursively. A synthesis of level *j* are shown in Equation 4.

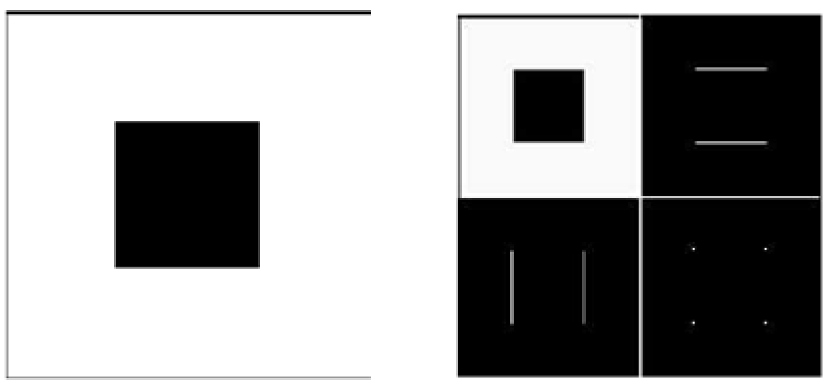
$$f(x) = \sum_k C_{0,k}\,\phi_{0,k}(x) = \sum_k C_{j+1,k}\,\phi_{j+1,k}(k) + \sum_{j=0}^{J} d_{j+1,k}\,\psi_{j+1,k}(x) \tag{4}$$

In two dimensional wavelet transform, initial image in each level, divides to four parts. The first part is an approximation of initial image and the other parts include components with high frequencies of image and show edge points in vertical, horizontal and diagonal directions. Figure1-(a) and (b) show this partitioning for the first and the second order of DWT on a 256 X 256 pixels image. After applying the first order of DWT on image, it is divided to four images I_{HH}, I_{LL}, I_{LH} and I_{HL} where H and L are high and low bounds respectively [11,12]. This operation can be performed For each of these four images repeatedly that the result is shown in part (b). Figure 2 shows a real example and the result of applying the first order of DWT on it in part (a) and (b) respectively.



(a) (b)

FIGURE 1: decomposition of an 256 x 256 pixels Image for the (a) first order, (b) second order of DWT



(a) (b)

FIGURE 2: (a): an image, (b): the result of applying the first order of DWT on (a)

Wavelet transform evaluates the difference of information in two different resolutions and creates a multi-resolution view of image's signal [11], so these transformation are useful for fabric defect detection.

3. GENETIC ALGORITHM

Genetic Algorithm is a stochastic optimization method. One of the main characteristic of this algorithm is that it leaves a set of best solutions in population. This algorithm has some operators for selecting the best solutions, crossing over and mutation in each generation [13]. GA starts with an initial and random population of solutions. Then a function that called fitness function, are calculated for each solution. Among initial population, solutions that their fitness values are higher than the others, are selected for next generation. In the other hand, Crossover and mutation operators add new solutions to next generation. this operation performs repeatedly to at the final, the algorithm converges to a good solution. In this paper, we used genetic algorithm for finding a suitable subset of wavelet coefficients of a special fabric textile.

The simplest form of GA involves three types of operators: [14]

- Selection: This operator selects chromosomes in the population for reproduction. The fitter the chromosome, the more times it is likely to be selected to reproduce.

- Crossover: This operator exchanges subsequences of two chromosomes to create two offspring. For example, the strings 10000100 and 11111111 could be crossed over after the third locus in each to produce the two offspring 10011111 and 11100100. This operator roughly mimics biological recombination between two single-chromosome (haploid) organisms.
- Mutation: This operator randomly flips some bits in a chromosome. For example, the string 00000100 might be mutated in its second position to yield 01000100. Mutation can occur at each bit position in a string with some probability, usually very small (e.g., 0.001).

the simple GA works as follows: [14]

1. Start with a randomly generated population of N L-bit chromosomes (candidate solutions to a problem).
2. Calculate the fitness F(x) of each chromosome x in the population.
3. Repeat the following steps (a)-(c) until N offspring have been created:
   a. Select a pair of parent chromosomes from the current population, with the probability of selection being an increasing function of fitness. Selection is done with replacement," meaning that the same chromosome can be selected more than once to become a parent.
   b. With probability P_c (the crossover probability), cross over the pair at a randomly chosen point (chosen with uniform probability) to form two offspring. If no crossover takes place, form two offspring that are exact copies of their respective parents.
   c. Mutate the two offspring at each locus with probability P_m (the mutation probability), and place the resulting chromosomes in the new population.
4. Replace the current population with the new population.
5. If the stopping condition has not satisfied, go to step 2.

4. FABRIC DEFECT DETECTION

Our proposed method has two phases. the First phase includes extracting all wavelet coefficients from a perfect fabric image and the second one includes finding a suitable subset of all coefficients using genetic algorithm. This subset can delete main fabric of image and indicate defects of fabric textile. So after these two phases, we can apply the suitable subset to the other images and gain the available defects in each of them.

Jasper in [15] showed that if the optimization problem in Eq. 5 can be solved, we can find a suitable subset of wavelet coefficients.

$$\min \; J = C * P * P^T * C^T$$

Where

$$C_{j(c)} = \sum_{k=0}^{n-1} c_k c_{k-2j} = 0 \tag{5}$$

In this equation, *J* function is the normal form of high frequency components of image, *n* is the number of coefficients, *c* is a subset of wavelet coefficients as a solution and *p* is an array of image's pixel values. Our goal is to find a *c* subset that minimize *J* function. For this purpose, we can use genetic algorithm. In this algorithm, each solution *c* is a chromosome and different fitness functions can be used. In this paper we evaluated some fitness functions but sum of squares and entropy functions had the best results.

Using a wavelet coefficients subset as a filter that filters images in both horizontal and vertical directions, has not good results for real fabrics that aren't symmetric. So by using two separable filters for horizontal and vertical directions, we can obtain the better results. If W_1 and W_2 be horizontal and vertical filters respectively and U be a perfect image, then the filtered image F, can be shown as Equation 6.

$$F = W_1 * U * W_2^T \tag{6}$$

Now we must calculate a fitness function for filtered image *F.* as mentioned previously, in this paper, we used sum of squares and entropy functions. These functions can be shown in Eq. 7 and 8 respectively.

$$Sum = \sum_{i=1}^{n} \sum_{j=1}^{m} F(i,j)^2 \tag{7}$$

$$\mathrm{H} = -\sum_{i=1}^{n} \mathrm{d(i)} * \log_2 (\mathrm{d(i)}) \tag{8}$$

In Eq. 7, n and m are the number of lines and columns in image and $F(i , j)$ is the pixel value of image, in line i and column j. in Eq. 8, n is the number of gray levels that here is 256 and $d(i)$ is probability of gray level i in filtered image F. this probability is a real number between 0 and 1. For example if 1024 pixels among all pixels have the same gray level value like 100, then can be written: d(100) = 1/64.

With these fitness functions in Eq. 7 and 8, the optimization problem that were shown in Eq. 5, can be written in form of Eq. 9 and 10 respectively. As mentioned previously, our purpose is to minimize J function.

$$J = \sum_{i=1}^{n} \sum_{j=1}^{m} (W_1 * U * W_2^T)^2 \tag{9}$$

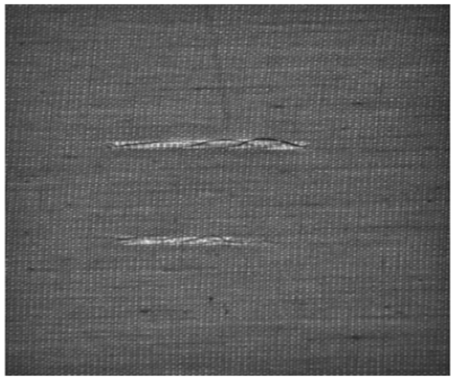
$$J = Entropy(W_1 * U * W_2^T) \tag{10}$$

The entropy function optimizes wavelet coefficients based on monotony of background. So if the entropy function value be small, the image will be more monotonic. But the sum of squares function optimizes image by minimizing its pixel values. furthermore entropy function has simpler computation in comparison with sum of squares. So we continued our experiments with entropy function as fitness function of GA.

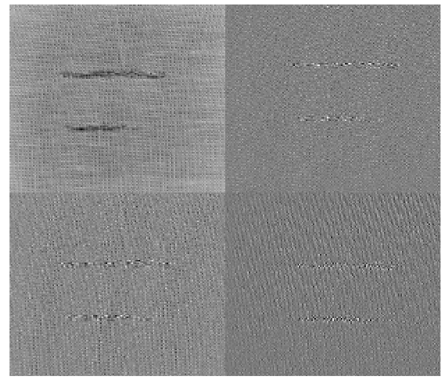
5. EXPERIMENTAL RESULTS

In this paper, for evaluating the proposed technique, we used two image databases that are called Tilda [16] and Brodatz [17]. Tilda includes 8 groups of different fabric textile types and Each group has 50 images with 512 x 768 pixels. Brodatz includes 112 types of fabric images with 640 x 640 pixels.

Figure 3 shows a defective fabric textile of Tilda database and the filtered image in part (a) and (b) respectively. The wavelet coefficients of each quarter was optimized by GA while the fitness function of GA was entropy. This Figure shows that the entropy function has good quality, because the background is monotonic and defects are appeared clearly.



(a)



(b)

FIGURE 3 : (a) a defective fabric textile, (b) filtered image by suitable coefficients for each quarter

In this experiment, first a perfect image from each fabric used for training phase and a suitable subset of its wavelet coefficients for each quarter was obtained. Figure 4 shows the histogram of quarter 4 of Figure 3-(b). this Figure shows a normal distribution. Most of pixels that are related to background are placed in the middle of distribution but defects are shown in the end of each littoral. So by selecting two thresholds in proximity of these two littoral, we can perform a thresholding process on the image and obtain a clear image for place detection of defects. The image can be more clear by applying some denoising and normalization processes on it.

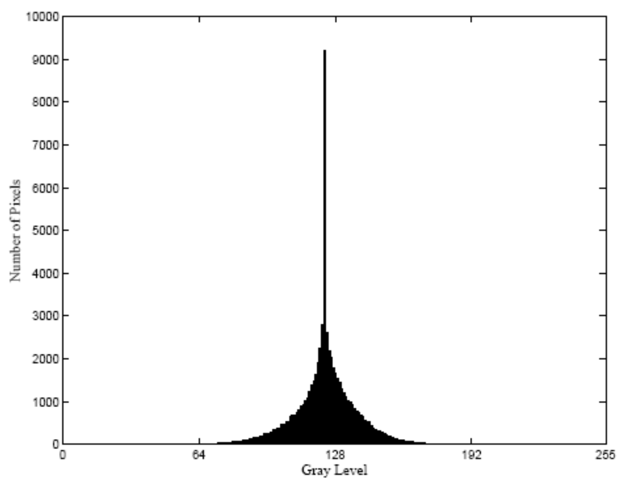


FIGURE 4: histogram of quarter 4 of Figure 3-(b)

Figure 5-(a) is a fabric textile with diagonal defects. Part (b) of this Figure shows quarter 4 of the filtered image that diagonal defects was appeared in it. Part (c) and (d) show the result of applying thresholding and denoising processes on part (b) and (c) respectively. Now part (d) of Figure 5 has suitable quality for defect detection.

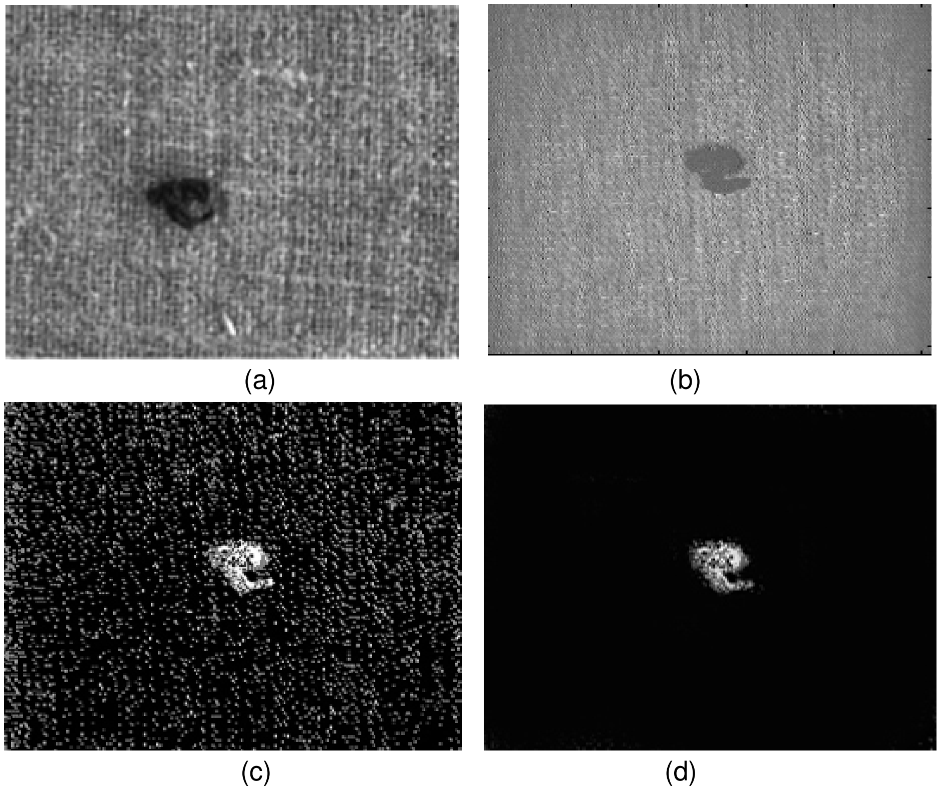


FIGURE 5: (a) a fabric with diagonal defects, (b) the quarter 4 of filtered image, (c) the result of applying thresholding process on part (b), (d) the result of applying denoising process on part (c)

Figure 6-(a) is a fabric textile with horizontal defects. Part (b) of this Figure shows quarter 3 of the filtered image that horizontal defects was appeared in it. Part (c) shows the result of applying thresholding on part (b), but this result is not suitable for denoising because the size of noise and defect areas are same nearly. So in this example we applied sliding window method on image in part (b) of Figure 6. This method builds 2 windows with size 17×17 and 9×9 for each pixel of image, Then calculates the standard deviation of each window and assignes the result of Equation (11) to central pixel as new value of it. In this Equation T is a suitable threshold.

$$\text{New value of Central pixel} = \begin{cases} if \quad \frac{\text{Standard Deviation of small window}}{\text{Standard Deviation of large window}} < T & 1 \\ otherwis & 0 \end{cases} \tag{11}$$

Part (d) of Figure 6 shows this result. In this image, size of defect areas are very larger than size of noise areas. So denoising process can be done on part (d). the result of denoising process is shown in part (e). Now part (e) of Figure 6 has suitable quality for defect detection.

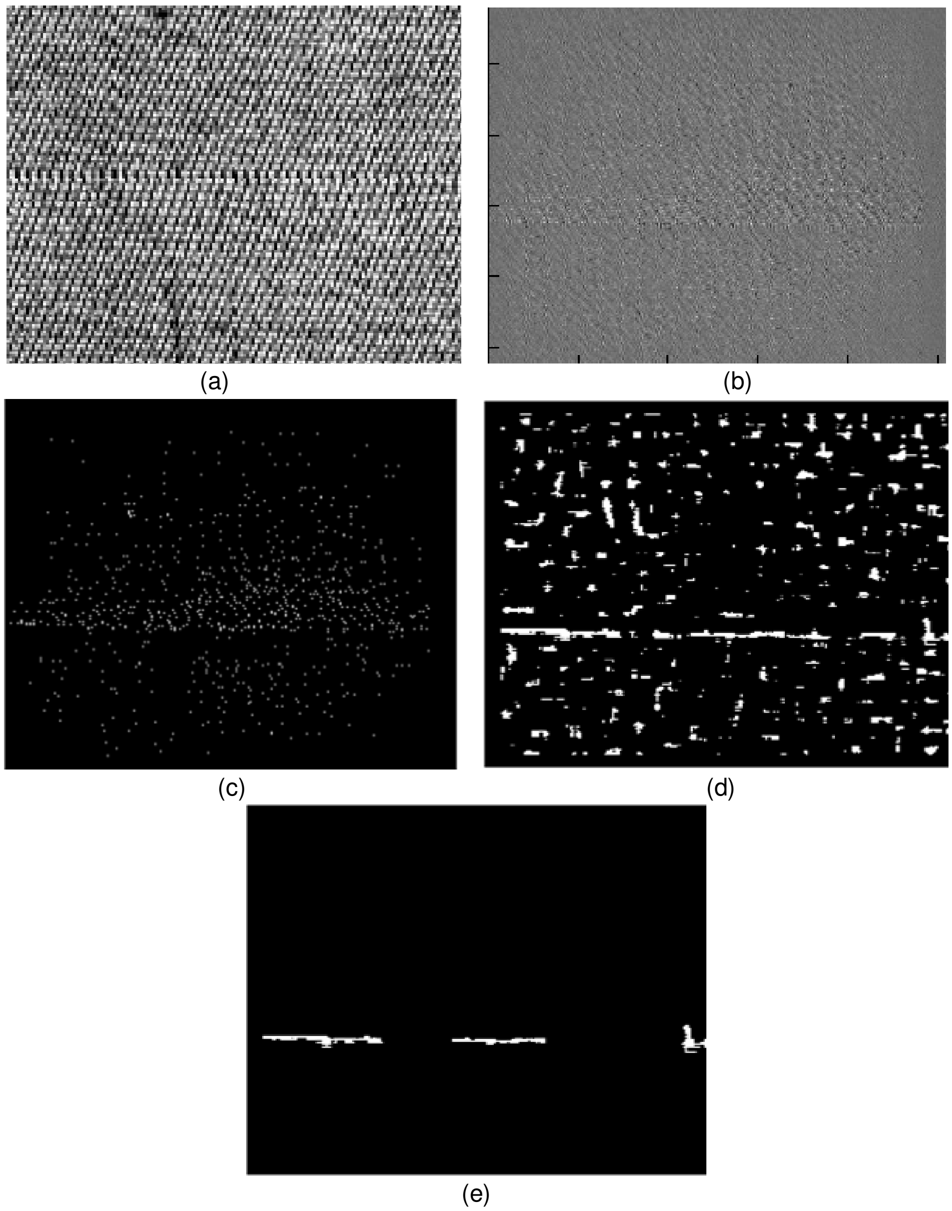


FIGURE 6: (a) a fabric with horizontal defects, (b) the quarter 3 of filtered image, (c) the result of applying thresholding process on part (b), (d) the result of applying sliding window method on part (b), (e) the result of applying denoising process on part (d)

In this experiment, first a perfect image from each fabric used for training phase and a suitable subset of its wavelet coefficients for each quarter, was obtained. Then in the test phase, the other perfect or defective images evaluated by our proposed technique. Table 1 and 2 show these results for Tilda and Brodatz databases, respectively. Since genetic algorithm starts with a random initial population, solutions that are produced by each population, must be same or close to each other. This characteristic is called repetition's capability. We tested this capability and its results were shown in Table3. In this table, for each quarter of two different fabric types, mean and standard deviation of ten runs with different initial populations were shown.

Fabric Type	Image number	Defective image number	Perfect image number	Detected defective image number	Detected perfect image number	Accuracy of correct detection
C1r1	18	12	6	11	6	94.4%
C1r2	22	14	8	14	7	95.4%
C2r3	20	16	4	15	4	95%

TABLE 1 : Fabric defect detection results for Tilda database

Fabric Type	Image number	Defective image number	Perfect image number	Detected defective image number	Detected perfect image number	Accuracy of correct detection
D77	100	23	77	19	76	95%
D79	100	27	73	23	70	93%
D105	100	56	44	51	39	90%

TABLE 2 : Fabric defect detection results for Brodatz database

Fabric type	Quarter number	Mean	Standard deviation
C1r1e3	2	5.9208	0.1092
C1r1e3	3	6.0158	0.1286
C1r1e3	4	5.9466	0.1297
C2r2e2	2	5.7401	0.1607
C2r2e2	3	5.8144	0.1707
C2r2e2	4	5.7530	0.1356

TABLE 3 : results of repetition's capability

6. CONCLUSION

As mentioned previously, in this paper we used a novel method for fabric textile defect detection. This process had two phases. the First phase includes extracting all wavelet coefficients from a perfect fabric image and the second one includes finding a suitable subset of all coefficients using genetic algorithm. The experimental results showed that our method have good accuracy for defect detection.

REFERENCES

1. K. Srinivasan and P.H. Dastor and P. Radhakrishnaihan, and S. Jayaraman "*FDAS: A Knowledge-based Frame Detection Work for Analysis of Defects in Woven Textile Structures*", Journal of Textile Institute, vol. 83, pp. 431-447, (1992).

2. R. Chin, "*Automated Visual Inspection Techniques and Applications: A Bibliography*", Pattern Recognition, 15(4): pp. 343–357, (1982).

3. Z. Guanng and W. Jianxia, " *Fabric Defect Detection Method Based on Image Distance Difference"*, Electronic Measurement and Instruments, (2007), pp. 822 -825.

4. D. Chetverikov and A. Hanbury, "Finding Defects in Texture Using Regularity and Local Orientation " , Pattern Recognition. 35(10), pp. 2165–2180, (2002).

5. X.Z. Yang and G. Pang and N. Yung, " Discriminative Fabric Defect Detection Using Adaptive Wavelets", Optical Engineering, pp. 3116-3126, December (2002).

6. X.Z. Yang and G. Pang and N. Yung, " Robust Fabric Dfect Detection and Classification Using Multiple Adaptive Wavelets", Image and Signal Processing, IEE Proceedings, volume 152, PP. 712-723 , (2005)

7. D. Chetverikov " Measuring the Degree of Texture Regularity", In Proceedings of International Conference on Pattern Recognition, 1984, volume 1, PP. 80–82, (1984).

8. Yu.zhang,zhaoyang Lu,Jing Li,"Fabric Defect Detection & classification using Gabor filters & Gaussian Mixture Model",Springer-LNCS,pp:635-644,(2010)

9. HAN leil,LI zhong,"quick defect detection based on structure character of woven fabric image & wavelet transform",computer engineering & design",(2009)

10. Rong Fu,Meihong Shi,Hongli Wei,Huijuan chen,"fabric defect detection based on adaptive local binary patterns",international conference on robotics & bi omimetics,pp:1336-1340,(2009)

11. S. Mallat," A Wavelet Tour of Signal Processing", Academic Press, 2nd ed., San Diego, (1999)

12. S. Mallat,"A Theory for Multiresolution Signal Decomposition: the Wavelet Representation", IEEE Transactions on Pattern Anal. and Mach. Intell., vol. 11, no. 7, pp. 674-693, (1989)

13. Z. Michalewicz, "Genetic Algorithms + Data Structures = Evolution Programs", AI Series. Springer-Verlag, New York, 3rd edition, (1996)

14. M. Mitchell, "Genetic Algorithm: An Overview", pp: 31-39, (1995).

15. L.W. Leung and B. King an, "Comparison of Image Data Fusion Techniques Using Entropy and INI" . 22nd Asian Conference on Remote Sensing, November (2001)

16. Workgroup on"Texture Analysis of DFG, "TILDA Textile Texture Database" , http://lmb.informatik.uni-freiburg.de/ research/dfg-texture/tilde

17. P.Brodatz, Textures: A Photographic Album for Artists and Designers , Dover, New York, (1966)

Steganalysis of LSB Embedded Images Using Gray Level Co-Occurrence Matrix

H.B.Kekre hbkekre@yahoo.com
Department of Computer Science
Mukesh Patel School of Technology Management
& Engineering, NMIMS University
Mumbai, India

A.A. Athawale athawalearchana@gmail.com
Department of Computer Science
Mukesh Patel School of Technology Management
& Engineering, NMIMS University
Mumbai, India

S.A.Patki sayli_patki@rediffmail.com
Department of Information Technology
K.J.Somaiya Institute of Engineering
& Information Technology, Mumbai University
Mumbai, India

Abstract

This paper proposes a steganalysis technique for both grayscale and color images. It uses the feature vectors derived from gray level co-occurrence matrix (GLCM) in spatial domain, which is sensitive to data embedding process. This GLCM matrix is derived from an image. Several combinations of diagonal elements of GLCM are considered as features. There is difference between the features of stego and non-stego images and this characteristic is used for steganalysis. Distance measures like Absolute distance and Euclidean distance are used for classification. Experimental results demonstrate that the proposed scheme outperforms the existing steganalysis techniques in attacking LSB steganographic schemes applied to spatial domain.

Keywords: Steganography, Steganalysis, LSB Embedding, GLCM, Distance Measures.

1. INTRODUCTION

Steganography is the art of passing information through apparently innocent files in a manner that the very existence of the message is unknown. The term steganography in Greek literally means, "Covered Writing" [1]. It uses the digital media such as text, image, audio, video and multimedia as a carrier (cover) for hiding private information in such a way that the third party cannot detect or even notice the presence of the communication. This gives indications that steganography can be used in criminal activities. The messages such as images, videos, sound files, text and other computer files can be hidden inside images or other digital objects which remains invisible to an ordinary observer. By embedding secret data into cover object, a stego object is obtained [2]. Steganalysis is the art of discovering and rendering useless the covert messages, hence breaking steganography. A steganalysis detector attempts to detect the presence or absence of an embedded message when presented with a stego signal. The basic rationale of steganalysis is that there should be differences between an original cover medium and its stego versions. Although the presence of embedded messages is often imperceptible to human eye, it may disturb the statistics of an image. Discovering the difference of some statistical characteristics between the cover and stego media becomes key issue in steganalysis [3].

In other view steganalysis techniques can be broadly divided as, (a). Passive steganalysis: Detect the presence or absence of a secret message in an observed media and (b). Active steganalysis: Extract an approximate version of the secret message or estimate some parameters such as embedding key, message length, etc. using a stego media [4].

2. RELATED WORK

In spatial domain, LSB-based steganography, in which the lowest bit plane of a bitmap image is used to convey the secret data, has long been used by steganographers. This is because the eye cannot detect the very small perturbations it introduces into an image and also because it is extremely simple to implement [5]. The tools used in this group include StegoDos, S – Tools, MandelSteg, Ezstego, Hide and Seek, Steganos [6] etc. LSB steganography methods can be divided into two classes, the LSB substitution [7] and LSB matching [7]. Several techniques for the steganalysis of the images for LSB embedding are present.

Fridrich J and Long M [8] proposed an algorithm for stego only attack. They analyzed the steganographic technique for the LSB embedding in 24-bit color images. The method is based on statistical analysis of the image colors in the RGB cube. Pfitzmann and Westfeld [9] introduced a method based on statistical analysis of Pairs of Values (PoVs) that are exchanged during message embedding. This method, which is the chi-square attack, is quite general and can be applied to many embedding paradigms besides the LSB embedding. Fridrich et al. [10] developed a steganographic method for detecting LSB embedding in 24-bit color images-the Raw Quick Pairs (RQP) method. This method is based on analyzing close pairs of colors created by LSB embedding. It works well if the number of unique colors in the cover image are less than 30 percent that of the total pixels. Sorina et .al [11], have introduced statistical sample pair approach to detect LSB steganography in digital signals such as images and audio. A quantitative steganalysis method to detect hidden information embedded by flipping pixels along the boundaries in binary images is presented in [12]. M. Abolghasemi et .al in [13] have proposed a method for detection of LSB data hiding based on Gray Level Co- Occurrence Matrix (GLCM). In [14] K.B.Raja et .al have proposed a method for LSB steganalysis to detect the embedded message length using pixel pair threshold. S. Mitra et.al [15], have described a detection theory based on statistical analysis of pixel pairs using their RGB components to detect the presence of hidden message in LSB steganography. They have used a fixed threshold method that resulted in poor detection rates. K.B.Raja et al [2], explain a LSB steganalysis method CPAVT based on variable threshold color pair analysis. They have employed "Color Density" as the measure to derive the variable threshold. S.Geetha et al [3], proposed another steganalysis method CCPASST based on variable threshold. Structural Similarity Index Measure is the measure used to obtain variable threshold.

3. GLCM (Gray Level Co-occurrence Matrix) and FEATURE EXTRACTION

A co-occurrence matrix is also referred to as co-occurrence distribution. It is defined over an image to be the distribution of co-occuring values at a given offset. Mathematically, a co-occurrence matrix C defined over an n x m image I, parametrized by an offset (Δx, Δy) is given as [13]:

$$C(i,j)=\sum_{p=1}^{n}\sum_{q=1}^{m}\begin{cases}1, & \text{if } I(p,q)=i \text{ and } I(p+\Delta x,q+\Delta y)=j\\ 0, & \text{otherwise}\end{cases}$$

Four different directions are selected for gray level co-occurrence matrix calculation, i.e. θ = 0°, 45°, 90° and 135° respectively. Thus four gray level co-occurrence matrixes: G1, G2, G3, G4 are obtained from these four directions respectively. From these four matrices the resultant co-occurrence matrix is generated as [13]:

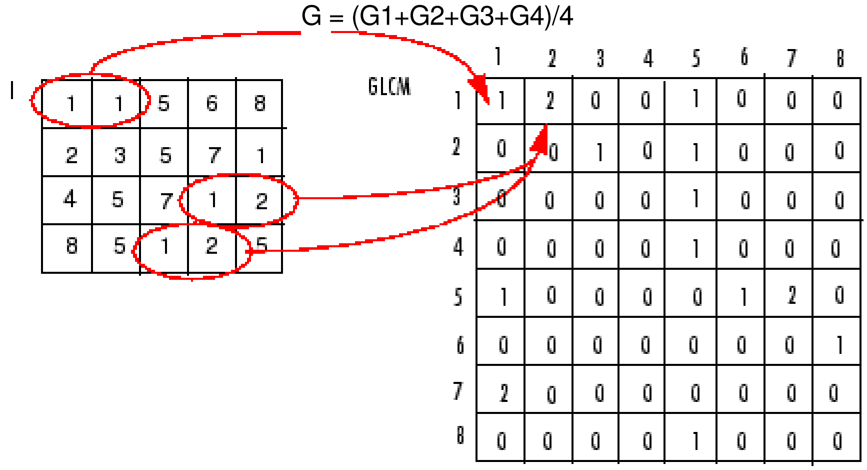


FIGURE 1: Matrix Elements and GLCM Matrix

The gray levels of neighboring pixels in natural images are often correlated, so the gray level co-occurrence matrix of the natural image tends to be diagonally distributed. However after data embedding the high concentration along the main diagonal of the matrix spreads as the high correlation between the pixels in the original image have been reduced as shown in Figure 1 and Figure 2.

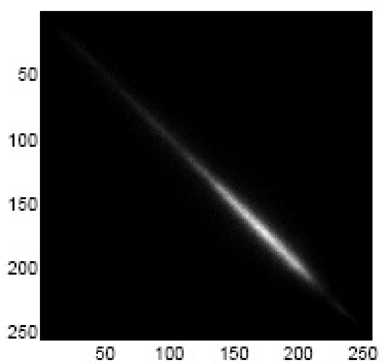


FIGURE 2: GLCM of Cover Image

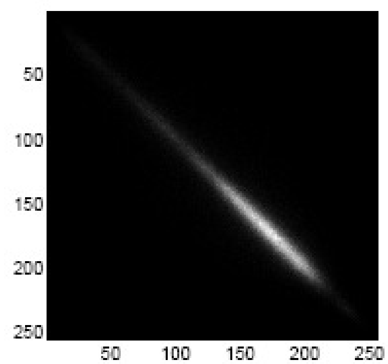


FIGURE 3: GLCM of Stego Image

Considering this asymmetry of the co-occurrence matrix, elements of the main diagonal (d0) and part of the upper (du1, du2) and lower (dl1, dl2) of main diagonal from GLCM are used to construct the feature vector [13]. Table 1 lists the 31 feature vectors used for the purpose of experiments. The 31 feature vectors are formed by considering the powerset of the five diagonals i.e. du2, du1, d0, dl1 and dl2

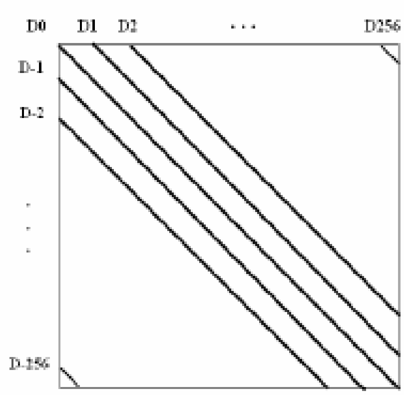


FIGURE 4: Diagonals of Co-occurrence Matrix as Features

F1(dl2,dl1,d0,du1,du2)	F17(d0,du2)
F2(dl1)	F18(dl2,du2)
F3(d0)	F19(dl1,d0,du1,du2)
F4(du1)	F20(dl2,dl1,d0,du1)
F5(du2)	F21(dl2,d0,du2)
F6(dl2,dl1)	F22(dl2,du1)
F7(du1,du2)	F23(dl1,du2)
F8(dl2,dl1,d0)	F24(dl2,dl1,d0,du2)
F9(d0,du1,du2)	F25(dl2,d0,du1,du2)
F10(dl1,d0,du1)	F26(dl2,du1,du2)
F11(dl2)	F27(dl1,du1,du2)
F12(dl1,du1)	F28(dl2,dl1,du1)
F13(dl2,dl1,du1,du2)	F29(dl2,dl1,du2)
F14(dl2,d0)	F30(dl2,d0,du1)
F15(dl1,d0)	F31(dl1,d0,du2)
F16(d0,du1)	

TABLE 1: List of Feature Vectors

Classification of images as stego or cover is done using two different distance measures: Absolute distance and Euclidean distance.

1. Feature vector of the test image is generated
2. Feature vector of the test image is compared with the feature vectors of all the images in the training database. Absolute distance measure and Euclidean distance measure is used to check the closeness of the test image and the training database images.
3. Distance values are sorted in ascending order and minimum of the values is considered
4. A threshold value is set to determine whether the image is stego image.

4. EXPERIMENTAL RESULTS

To test the performance a database of BMP images is used. It consists of 30 color images and 30 grayscale images of size 128 x 128. This database is augmented with the stego versions of these images using LSB steganography. Different payload strengths were used i.e. 25%, 45%, 50%, 90%, 100% of the size of the cover image. So the database consists of 180 color images (cover and stego) and 180 gray scale images (cover and stego). In the experiments, 30 randomly selected images (cover and stego) are taken as training images.

After prolonged testing with database of images, threshold is selected on trial and error basis. Using this threshold, stego images are identified from the database. We have considered four different threshold values 100, 150, 200 and 250. With increase in the threshold the results improve, so we have considered the results with maximum threshold i.e. 250. In case of grayscale images operations such as obtaining the GLCM and extracting the features from the same are performed on the image as a whole. On the other hand color images are first separated in to three planes (Red, Green and Blue), and operations of obtaining the GLCM and extracting the features from the same are performed on each of the planes separately. The maximum of the results of the three planes are taken into consideration.

Table 2 shows the percentage detection for 31 features using Absolute distance in grayscale images and Table 3 shows the percentage detection for 31 features using Absolute distance in color images. Table 4 shows the percentage detection for 31 features using Euclidean distance in grayscale images and Table 5 shows the percentage detection for 31 features using Euclidean distance in color images. Percentage detection indicates out of 30 test stego images, the number of images that are detected as stego.

Feature	Length of Embedding				
	25%	45%	50%	90%	100%
F1	23	37	23	17	23
F2	100	100	97	87	90
F3	70	73	67	47	50
F4	93	97	90	87	87
F5	83	90	80	70	77
F6	77	77	73	53	60
F7	77	77	73	50	60
F8	43	53	47	20	27
F9	40	50	47	20	27
F10	43	50	43	23	30
F11	83	87	77	70	83
F12	83	83	77	60	73
F13	47	53	50	20	27
F14	63	67	60	30	37
F15	63	67	63	33	37
F16	63	67	63	30	37
F17	63	67	57	30	37
F18	70	67	67	47	57
F19	30	40	27	20	27
F20	30	40	23	20	27
F21	37	50	50	20	27
F22	73	77	73	53	60
F23	80	77	73	50	60
F24	30	40	23	20	27
F25	30	40	23	20	27
F26	67	63	60	20	33
F27	67	67	63	27	43
F28	67	67	67	27	43
F29	63	67	60	20	37
F30	43	53	47	20	27
F31	40	50	47	20	27

TABLE 2: Detection accuracy comparison for 31 features: Absolute Distance and Grayscale images

Feature	Length of Embedding				
	25%	45%	50%	90%	100%
F1	33	33	33	20	20
F2	100	100	100	97	97
F3	80	83	87	63	57
F4	100	100	100	97	93
F5	93	90	93	87	87
F6	77	83	90	73	63
F7	77	83	90	73	63
F8	53	53	50	33	33
F9	47	50	50	33	33
F10	50	53	53	33	37
F11	90	90	93	90	87
F12	93	90	97	80	80
F13	50	53	53	43	30
F14	60	67	70	53	40
F15	63	70	83	53	43
F16	63	70	77	53	43
F17	60	70	67	47	40
F18	73	80	77	70	60
F19	40	43	37	27	30
F20	40	43	40	27	27
F21	47	50	53	30	33
F22	73	83	90	73	63
F23	77	83	90	73	63
F24	40	40	37	23	27
F25	37	40	37	23	27
F26	63	70	70	50	40
F27	70	73	80	53	40
F28	70	77	77	53	40
F29	63	70	70	47	40
F30	53	53	47	37	33
F31	47	53	53	33	33

TABLE 3: Detection accuracy comparison for 31 features: Absolute Distance and Color images

Feature	Length of Embedding				
	25%	45%	50%	90%	100%
F1	87	93	93	73	77
F2	100	100	100	93	93
F3	97	100	100	83	83
F4	100	100	100	93	93
F5	100	100	100	97	97
F6	100	100	100	87	93
F7	100	100	100	87	93
F8	87	97	97	80	77
F9	87	97	97	80	77
F10	87	100	100	80	83
F11	100	100	100	97	97
F12	100	100	100	90	93
F13	97	97	97	83	90
F14	87	97	97	80	80
F15	93	100	100	80	83
F16	93	100	100	80	83
F17	87	97	97	80	80
F18	100	100	100	93	97
F19	87	97	97	77	77
F20	87	97	97	73	77
F21	87	97	97	77	77
F22	100	100	100	87	93
F23	100	100	100	87	93
F24	87	97	97	73	77
F25	87	97	97	73	77
F26	97	97	97	83	90
F27	100	100	100	83	93
F28	100	100	100	83	93
F29	97	97	97	83	90
F30	87	97	97	77	77
F31	87	97	97	80	77

TABLE 4: Detection accuracy comparison for 31 features: Euclidean Distance and Grayscale images

Feature	Length of Embedding				
	25%	45%	50%	90%	100%
F1	90	93	97	80	73
F2	100	100	100	97	97
F3	90	97	100	87	83
F4	100	100	100	97	97
F5	100	100	100	100	97
F6	100	100	100	93	93
F7	100	100	100	93	93
F8	90	93	97	83	80
F9	90	93	97	83	80
F10	90	93	97	87	83
F11	100	100	100	100	97
F12	100	100	100	97	93
F13	97	97	97	90	90
F14	90	93	97	83	80
F15	90	93	97	87	83
F16	90	93	97	87	83
F17	90	93	97	83	80
F18	97	97	97	97	97
F19	90	93	97	80	80
F20	90	93	97	80	73
F21	90	93	97	83	80
F22	100	100	100	93	93
F23	100	100	100	93	93
F24	90	93	97	80	73
F25	90	93	97	80	73
F26	97	97	97	93	90
F27	100	97	97	93	93
F28	100	97	97	93	90
F29	97	97	97	93	90
F30	90	93	97	80	80
F31	90	93	97	83	80

TABLE 5: Detection accuracy comparison for 31 features: Euclidean Distance and Color images

5. CONSLUSION

This paper discusses a steganalysis method based on features that are extracted from co-occurrence matrix of an image. Two different distance measures: Absolute and Euclidean are used for the purpose of classification. This scheme outperforms previous works in steganalysis for LSB hiding. It works in case of both grayscale and color images. Euclidean distance gives the best results. It is observed that results obtained using Euclidean distances are better than Absolute distance by around 329% in grayscale images and by 265% in color images. Detection accuracy in case of color images is better than that of grayscale images by around 18% in Absolute distance and almost same in Euclidean distance. Superiority is observed for low embedding rates. The feature vectors which consist of the diagonal d0 exhibit poor results as compared to feature vectors that do not contain the diagonal d0.

6. REFERENCES

1. S. K. Jena, G.V.V. Krishna, "Blind Steganalysis: Estimation of Hidden Message Length", *International Journal of Computers, Communications and Control*, Vol. II, No. 2: 149-158

2. K. B. Raja, Shankara N, Venugopal K. R., and L.M. Patnaik, "Steganalysis of LSB Embedded Images using Variable Threshold Color Pair Analysis", *International Journal of Information Processing*, Vol. 1, 2007

3. S.Geetha, Siva S. Sivatha sindhu, R. Renganathan, P. Janaki Raman and Dr. N. Kamraj, "StegoHunter: Steganalysis of LSB Embedded Images based on Stego – Sensitive Threshold Close Color Pair Signature", *Sixth Indian Conference on Computer Vision, Graphics and Image Processing*, 2008, pp. 281 – 288, 2008

4. M.Abolghasemi, H.aghainia, K.Faez, M.A.Mehrabi, "LSB data hiding detection based on gray level co-occurrence matrix", *International symposium on Telecommunications,*2008, pp. 656-659, 2008

5. J. Fridrich, M. Goljan, r. Du, "Reliable detection of LSB steganography in grayscale and color images", *Proceeding of ACM, Special Session on Multimedia Security and Watermarking*, Ottawa, Canada, pp. 27-30, 2001.

6. Wayner P., "Disappearing Cryptography", *Morgan Kaufmann Publisher*, 2002

7. S. Manoharan, "An Empirical Analysis of RS Steganalysis", *International Conference on Internet Monitoring and Protection,* 2008, pp. 172-177.

8. Westfeld and A. Piftzmann, "Attacks on Steganographic Systems", Proc. 3rd International Workshop on Information Hiding*, LNCS 1768,* pp. 61-76, Springer-Verlag, 2000

9. J. Fridrich, M. Goljan and R. Du, "Detecting LSB steganography in color and gray-scale images", *IEEE Multimedia,* vol. 8, no. 4, pp. 22-28,2001

10. S. Dumitrescu, X. Wu, and Z. Wang, "Detection of LSBsteganography via sample pair analysis", *IEEE Trans. SignalProcessing,* vol. 51, no. 7, pp. 1995-2007, 2003

11. Q. Liu, A. H.Sung, Jianyun Xu and Bernardete M. Riberio, "Image Complexity and Feature Extraction for Steganalysis of LSB Matching Steganography", *International Conference on Pattern Recognition*, 2006,

12. M.Abolghasemi, H.aghainia, K.Faez, M.A.Mehrabi, "steganalysis of LSB Matching based on co-occurrence matrix and removing most significant planes", *International conference on Intelligent Information Hiding and Multimedia Signal Processing,*2008, pp. 1527- 1530, 2008

13. A. D. Ker, "Steganalysis of LSB Matching in Grayscale Images",*IEEE Signal Processing Letters*, Vol. 12, No. 6, june 2005

14. M.A.Mehrabi, H. Aghaeinia, M. Abolghasemi, "Steganalysis of LSB Matching Steganography by Removing Most Significant Bit Planes", *International Symposium on Telecommunications*, 2008, pp. 731-734, 2008

15. F. Huang, B. Li, J. Huang, “Attack LSB Matching Steganography by counting alteration rate of the numnber of neighbourhood gray levels”, *International Conferenece on Image Processing*, 2007, pp. 401 – 404, 2007

Unconstrained Optimization Method to Design Two Channel Quadrature Mirror Filter Banks for Image Coding

Anamika Jain ajain_bpl@yahoo.in
Department of Electronics and
Communication Engineering
M.A.N.I.T., Bhopal, 462051, INDIA

Aditya Goel adityagoel2@rediffmail.com
Department of Electronics and
Communication Engineering
M.A.N.I.T., Bhopal, 462051, INDIA

Abstract

This paper proposes an efficient method for the design of two-channel, quadrature mirror filter (QMF) bank for subband image coding. The choice of filter bank is important as it affects image quality as well as system design complexity. The design problem is formulated as weighted sum of reconstruction error in time domain and passband and stop-band energy of the low-pass analysis filter of the filter bank .The objective function is minimized directly, using nonlinear unconstrained method. Experimental results of the method on images show that the performance of the proposed method is better than that of the already existing methods. The impact of some filter characteristics, such as stopband attenuation, stopband edge, and filter length on the performance of the reconstructed images is also investigated.

Keywords: Sub-Band Coding, MSE (mean square error), Perfect Reconstruction, PSNR(Peak Signal to Noise Ratio); Quadrature Mirror filter).

1. INTRODUCTION

Quadrature mirror filter (QMF) banks have been widely used in signal processing fields, such as sub-band coding of speech and image signals [1–4], speech and image compression [5,6], transmultiplexers, equalization of wireless communication channels, source coding for audio and video signals , design of wavelet bases [7], sub-band acoustic echo cancellation , and discrete multitone modulation systems. In the design of QMF banks, it is required that the perfect, reconstruction condition be achieved and the intra-band aliasing be eliminated or minimized. Design methods [8,9] developed so far involve minimizing an error function directly in the frequency domain or time domain to achieve the design requirements. In the conventional QMF design techniques [10]-[19] to get minimum point analytically, the objective function, is evaluated by discretization, or iterative least squares methods are used which are based on the linearization of the error function to, modify the objective function. Thus, the performance of the QMF bank designed degrades as the solution obtained is the minimization of the discretized version of the objective function rather than the objective function itself, or computational complexity increased.

In this paper a nonlinear optimization method is proposed for the design of two-channel QMF bank. The perfect reconstruction condition is formulated in the time domain to reduce computation complexity and the objective function is evaluated directly [12-19].Various design techniques including optimization based [20], and non optimization based techniques have been reported in literature for the design of QMF bank. In optimization based technique, the design problem is formulated either as multi-objective or single objective nonlinear optimization problem, which is solved by various existing methods such as least square technique, weighted least square (WLS) technique [14-17] and genetic algorithm [21]. In early stage of research, the design methods developed were based on direct minimization of error function in frequency domain [8]. But due to high degree of nonlinearity and complex optimization technique, these methods were not suitable for the filter with larger taps. Therefore, Jain and Crochiere [9] have introduced the concept of iterative algorithm and formulated the design problem in quadratic form in time domain. Thereafter, several new iterative algorithms [10, 12-21] have been developed either in time domain or frequency domain. Unfortunately, these techniques are complicated, and are only applicable to the two-band QMF banks that have low orders. Xu *et al* [10,13,17] has proposed some iterative methods in which, the perfect reconstruction condition is formulated in time domain for reducing computational complexity in the design. For some application, it is required that the reconstruction error shows equiripple behaviour, and the stopband energies of filters are to be kept at minimum value. To solve these problems, a two-step approach for the design of two-channel filter banks was developed. But the approach results in nonlinear phase, and is not suitable for the wideband audio signal. Therefore, a modified method for the design of QMF banks using nonlinear optimization has developed in which prototype filter coefficients are optimized to minimize the combination of reconstruction error, passband and stopband and residual energy.

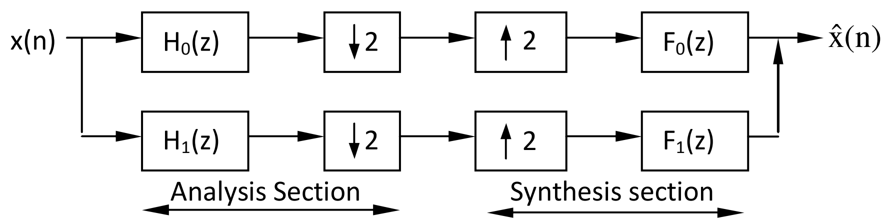


FIGURE1: Quadrature Mirror Filter Bank

A typical two-channel QMF bank shown in Figure 1, splits the input signal *x(n)* into two subband signals having equal band width, using the low-pass and high-pass analysis filters $H_0(z)$ and $H_1(z)$, respectively. These subband signals are down sampled by a factor of two to achieve signal compression or to reduce processing complexity. At the output end, the two subband signals are interpolated by a factor of two and passed through lowpass and highpass synthesis filters, $F_0(z)$ and $F_1(z)$, respectively. The outputs of the synthesis filters are combined to obtain the reconstructed signal $\hat{x}$(n). The reconstructed signal $\hat{x}(n)$ is different from the input signal *x (n)* due to three errors: aliasing distortion (ALD), amplitude distortion (AMD), and phase distortion (PHD). While designing filters for the QMF bank, the main stress of most of the researchers has been on the elimination or minimization of the three distortions to obtain a perfect reconstruction (PR) or nearly perfect reconstruction (NPR) system. In several design methods reported [17–23], aliasing has been cancelled completely by selecting the synthesis filters cleverly in terms of the analysis filters and the PHD has been eliminated using the linear phase FIR filters. The overall transfer function of such an alias and phase distortion free system turns out to be a function of the filter tap coefficients of the lowpass analysis filter only, as the highpass and lowpass analysis filters are related to each other by the mirror image symmetry condition around the quadrature frequency $\pi/2$. Therefore, the AMD can be minimized by optimizing the filter tap weights of the lowpass analysis filter. If the characteristics of the lowpass analysis filter are assumed to be ideal in its passband and stopband regions, the PR condition of the alias and phase distortion free QMF bank is automatically satisfied in these regions, but not in the transition band. The objective function to be minimized is a linear combination of the reconstruction error in time domain and passband and stopband residual energy of the lowpass analysis filter of the filter bank .A nonlinear unconstrained optimization method [20] has been used to minimize the objective function by optimizing the coefficients of the lowpass analysis filter. A comparison of the design results of the proposed method with that of the already existing methods shows that this method is very effective in designing the two channel QMF bank, and gives an improved performance.

The organization of the paper is as follows: in Section 2, a relevant brief analysis of the QMF bank is given. Section 3 describes the formulation of the design problem to obtain the objective function. A mathematical formulation to minimize the objective function by using unconstrained optimization method has been explained in Section 4 and Section 5 presents the proposed design algorithm. In Section 6, two design examples (cases) are presented to illustrate the proposed design algorithm. Finally, an application of the proposed method in the area of subband coding of images is explained. A comparison of the simulation results of the proposed algorithm with that of the already existing methods is also discussed.

2. ANALYSIS OF THE TWO-CHANNEL QMF BANK

The *z*-transform of the output signal $\hat{x}$ *(n)*, of the two channel QMF bank, can be written as [18–20, 23]

$$\hat{X}(z) = ½[H_0(z)F_0(z) + H_1(z)F_1(z)]X(z) + ½[H_0(-z)F_0(z) + H_1(-z)F_1(z)]X(-z). \quad (1)$$

Aliasing can be removed completely by defining the synthesis filters as given below [1, 20–23]

$$F_0(z) = H_1(-z) \text{ and } F_1(z) = -H_0(-z). \quad (2)$$

Therefore, using Eq. (2) and the relationship $H_1(z) = H_0(-z)$ between the mirror image filters, the expression for the alias free reconstructed signal can be written as

$$\hat{X}(z) = ½[H_0(z)H_1(-z) - H_1(z)H_0(-z)]X(z) = ½[H_0^2(z) - H_0^2(-z)]X(z) \quad (3)$$

or

$$\hat{X}(z) = T(z)\, X(z), \quad (4)$$

where *T (z)* is the overall system function of the alias free QMF bank, and is given by

$$T(z) = ½[H_0^2(z) - H_0^2(-z)] \quad (5)$$

To obtain perfect reconstruction, AMD (amplitude distortion) and PHD (phase distortion) should also be eliminated, which can be done if the reconstructed signal $\hat{x}$ *(n)* is simply made equal to a scaled and delayed version of the input signal *x(n)*. In that situation the overall system function, must be equal to:

$$T(z) = cz^{-(N-1)} \quad (6)$$

where (N-1) is reconstruction delay .The perfect reconstruction condition in time-domain can be expressed by using the convolution matrices as [20]

$$Bh_0 = m$$

$$B = [d_1 + d_N, d_2 + d_{N-1}, ..., d_{N/2} + d_{N/2+1}]$$

$$D = [d_1, d_2, ..., d_N]$$

$$= 2\begin{bmatrix} h_0(1) & h_0(0) & 0 & \cdots & 0 \\ h_0(3) & h_0(2) & h_0(1) & h_0(0) & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ h_0(N-1) & h_0(N-2) & h_0(N-3) & \cdots & h_0(0) \end{bmatrix}$$

$$h_0 = [h_0(0), h_0(1), ... h_0(N/2-1)]^T$$

$$m = [0, 0, ...1]^T \quad (7)$$

Where $h_0(n)$, for n = 0, 1, 2. . . N/2-1, is the impulse response of filter H_0. To satisfy the linear phase FIR property, the impulse response $h_0(n)$ of the lowpass analysis filter can be assumed to be symmetric. Therefore,

$$h_0(n) = \begin{cases} h_0(N-1-n).\ 0 \le n \le N-1 \\ 0, n < 0 \text{ and } n \ge N \end{cases} \quad (8)$$

For real $h_0(n)$, $H_R(\omega)$ amplitude function is an even function of ω. Hence, by substituting Eqn. (8) into Eqn. (5), the overall frequency response of the QMF bank can be written as:

$$T(e^{j\omega}) = (e^{-j\omega(N-1)}/2)[\left|H_0(e^{j\omega})\right|^2 - (-1)^{(N-1)}\left|H_0(e^{j(\pi-\omega)})\right|^2] \tag{9}$$

If the filter length, *N*, is odd, above equation gives $T(e^{j\omega}) = 0$ at $\omega = \pi/2$, implying severe amplitude distortion. In order to cancel the aliasing completely, the synthesis filters are related to the analysis filters by Eqn. (2) and $H_1(z) = H_0(-z)$. It means that the overall design task reduces to the determination of the filter tap coefficients of the linear phase FIR low-pass analysis filter $H_0(z)$ only, subject to the perfect reconstruction condition of Eqn. (7). Therefore, we propose to minimize the following objective function for the design of the QMF bank, by optimizing the filter tap weights of the lowpass filter $H_0(z)$

$$\Phi = \alpha_1 E_p + \alpha_2 E_s + \beta E_r \tag{10}$$

where $\alpha_1, \alpha_2, \beta$ are real constants, E_p, E_s are the measures of passband and stopband error of the filter $H_0(z)$, and E_r is the square error of the overall transfer function of the QMF bank in time domain, respectively.

The square error E_r is given by

$$E_r = (Bh_0 - m)^T (Bh_0 - m) \tag{11}$$

3. PROBLEM FORMULATION

3.1. PASS-BAND ERROR

For even *N*, the frequency response of the lowpass filter $H_0(z)$ is given by

$$H_0(e^{j\omega}) = e^{-j\omega(N-1)/2} \sum_{n=0}^{(N/2-1)} 2h_0(n)\cos(\omega(N-1)/2 - n) \tag{12}$$

$$= e^{-j\omega(N-1)/2} \sum_{n=0}^{(N/2-1)} b(n)\cos(\omega(N-1)/2 - n) \tag{13}$$

$$= e^{-j\omega(N-1)/2} H_R(\omega) \tag{14}$$

Where

$$b(n) = 2h_0(n) \tag{15}$$

and $H_R(\omega)$ is the magnitude function defined as

$$H_R(\omega) = \mathbf{b}^T\mathbf{c}. \tag{16}$$

Vectors **b** and **c** are

$$b = [b(0)\ b(1)\ b(2) \ldots b(N/2-1)]^T$$

$$c = [\cos\omega((N-1)/2)\ \ \cos\omega((N-1)/2-1)\ \ldots\ \cos(\omega/2)]^T$$

Mean square error in the passband may be taken as

$$E_p = (1/\pi)\int_0^{\omega_p} b^T(1-c)(1-c)^T b\, d\omega \tag{17}$$

$$= b^T Q b \tag{18}$$

where matrix Q is

$$Q = (1/\pi)\int_0^{\omega_p} (1-c)(1-c)^T d\omega \tag{19}$$

With (m, n) th element given by

$$q_{mn} = (1/\pi)\int_0^{\omega_p} [(1-\cos\omega(N-1)/2-m)(1-\cos\omega(N-1)/2-n)]d\omega \tag{20}$$

3.2 STOPBAND ERROR

Mean square error in the stopband may be taken as

$$E_s = (1/\pi)\int_{\omega_s}^{\pi} b^T cc^T b d\omega \tag{21}$$

$$= b^T P b \tag{22}$$

where matrix **P** is

$$P = (1/\pi)\int_{\omega_s}^{\pi} cc^T d\omega \tag{23}$$

With (m, n)th element given by

$$p_{mn} = (1/\pi)\int_{\omega_s}^{\pi} (\cos\omega(N-1)/2-m)(\cos\omega(N-1)/2-n)d\omega \tag{24}$$

3.3 THE RECONSTRUCTION SQUARE ERROR

The square error E_r is given by

$$E_r = (Bh_0 - m)^T (Bh_0 - m) \tag{25}$$

is used to approximate the prefect reconstruction condition in the time-domain in which **B** and m are all defined as in eqn.(7).

- **Minimization Of The Objective Function**

Using Eqs. (19), (23), and (26), the objective function Φ given by eqn. (11) can be written as

$$\Phi = \alpha_1 b^T Q b + \alpha_2 b^T P b + \beta E_r$$

$$\Phi = 4\alpha_1 h_0{}^T Q h_0 + 4\alpha_2 h_0{}^T P h_0 + \beta E_r \tag{26}$$

It is in quadratic form without any constraints. Therefore the design problem is reduced to unconstrained optimization of the objective function as given in eqn. (26).

4. THE DESIGN ALGORITHM

In the designs proposed by Jain–Crochiere [9], and Swaminathan–Vaidyanathan [26], the unit energy constraint on the filter coefficients was also imposed. In the algorithm presented here, the unit energy constraint is imposed within some prespecified limit. The design algorithm proceeds through the following steps:

(1) Assume initial values of $\alpha_1, \alpha_2, \beta$, ϵ_1, ω_p, ω_s, and *N*.

(2) Start with an initial vector $\mathbf{h}_0 = [h_0(0)\ h_0(1)\ h_0(2)\ ...\ h_0((N/2)-1)]^T$; satisfying the unit energy constraint within a prespecified tolerance, i.e.

$$u = \left|1 - 2\sum_{k=0}^{N/2-1} h_0^2(k)\right| < \epsilon_1 \tag{27}$$

(3) Set the function tolerance, convergence tolerance .
(4) Optimize objective function eqn. (26) using unconstrained optimization method for the specified tolerance.
(5) Evaluate all the component filters of QMF bank using h_0.

The performance of the proposed filter and filter bank is evaluated in terms of the following significant parameters:

Mean square error in the passband

$$E_p = (1/\pi)\int_0^{\omega_p} [|H_0(0)| - |H_0(f)|]^2 d\omega \tag{28}$$

Mean square error in the stopband

$$E_s = \int_{\omega_s}^{\pi} \left| H_0(\omega) \right|^2 d\omega \tag{29}$$

stopband edge attenuation

$$A_s = -20 \log_{10}(H_0(\omega_s)) \tag{30}$$

Measure of ripple

$$(\in) = \max_{\omega} \left| 10 \log_{10} \left| T(\omega) \right| \right| - \min_{\omega} \left| 10 \log_{10} \left| T(\omega) \right| \right| \tag{31}$$

5. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed technique for the design of QMF bank has been implemented in MATLAB. Two design cases are presented to illustrate the effectiveness of the proposed algorithm. The method starts by initializing value of the filter coefficients $h_0(n)$ to zero for all n, except that $h_0(N/2-1) = h_0(N/2) = 2^{-1/2}$, $0 \le n \le N-1$ and using half of its coefficients which **t**hen be solved using unconstrained optimization (interior reflective Newton method) problem. With this choice of the initial value of the filter coefficients, the unit energy constraint is satisfied.

In both designs, stopband first lobe attenuation (A_s) has been obtained and the constants $\alpha_1, \alpha_2, \beta$, ε_1 have been selected by trial and error method to obtain the best possible results. The parameters used in the two designs, which will be referred to as Cases 1 and 2, are N =24, α_1 =.01, α_2 =0.1, β =.00086, $\text{function}_{\text{tolerance}}$=1e-6, $X_{\text{tolerance}}$=1e-8, ε_1 = 0.15, (IT) =38 ω_p = 0.4π, ω_s = 0.6π, tau = 0.5, and ***N*** = 32, α_1 =.1, α_2 =0.1, β =.00086, $\text{function}_{\text{tolerance}}$=1e-6, $X_{\text{tolerance}}$=1e-8, ε_1 = 1e-8, ω_p = 0.4π, ω_s = 0.6π, tau = 0.5, respectively. For comparison purposes the method of Chen and Lee [8] was applied to design the QMF banks with parameters specified above and using the same initial **h**, as by the proposed method to both the design examples (cases) respectively. The comparisons are made in terms of phase response, passband energy, stopband energy, stopband attenuation and peak ripple $(\in)$.

Case 1

For N = 24, $\omega p = 0.4\pi$, $\omega s = 0.6\pi$, $\alpha 1$ = .1, $\alpha 2$ =.1, ε_1 = 0.15, β = 1, the following filter coefficients for $(0 \le n \le N/2-1)$ are obtained

h_0 (n) = [-0.0087 -0.0119 0.0094 0.0221 -0.0123 -0.0332 0.0235 0.0540 -0.0463 -0.0970 0.1356 0.4623]

The corresponding normalized magnitude plots of the analysis filters $H_0(z)$ and $H_1(z)$ are shown in Figure 2a & 2c. Figure 2e shows the reconstruction error of the QMF bank (in dB). The significant parameters obtained are: Ep = .1438, Es = 1.042×10−6, As = 4**5.831 dB** and $(\in)$ = **0.9655.**

Case 2

For N = 32, $\omega_p = 0.4\pi$, $\omega_s = 0.6\pi$, $\alpha_1 = 0.1$, $\alpha_2 = 0.1$, $\varepsilon_1 = 0.15$, $\beta = 0.00086$, the following filter coefficients for $(0 \le n \le N/2-1)$ are obtained

h_0 (n) = [-0.0034 -0.0061 0.0020 0.0104 -0.0021 -0.0154 0.0050 0.0237 -0.0102 -0.0349 0.0218 0.0549 -0.0460 -0.0987 0.1343 0.4628]

The corresponding normalized magnitude plots of the analysis filters H_0 *(z)* and H_1 *(z)* are shown in Fig. 2b & 2d. Figure 2f shows the reconstruction error of the QMF bank (in dB). The significant parameters obtained are: Ep = .0398, Es = 2.69×10−7, As = **53.391** dB, and $(\in)$ = **0.27325**.

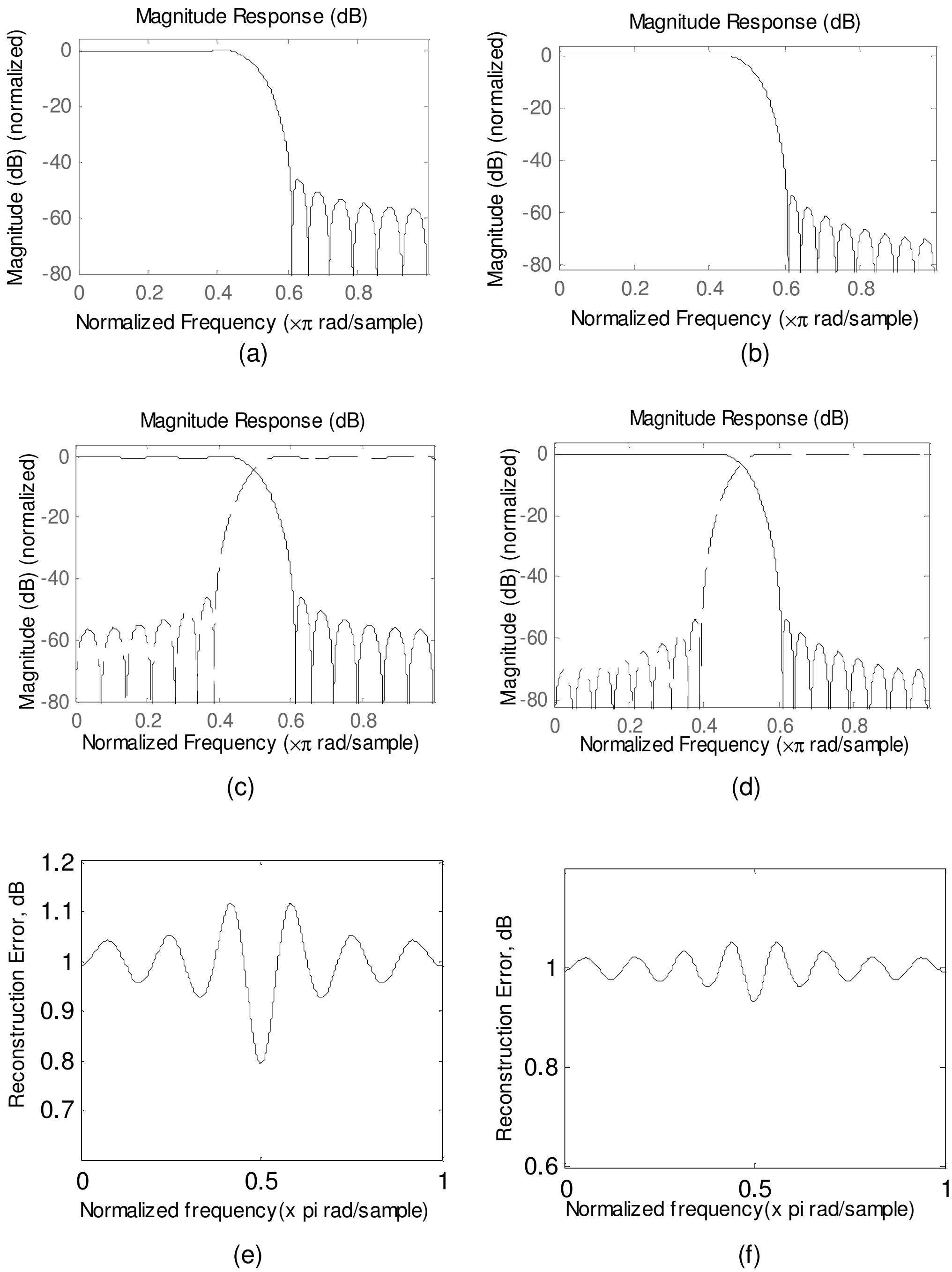


FIGURE 2: (a) Amplitude response of the prototype analysis filters for $N = 24$. (b) Amplitude response of the prototype analysis filters for N=32(c) Magnitude response of overall filter bank N=24 (d) Magnitude response of overall filter bank N=32 (e) Reconstruction error for overall filter bank (N=24) in dB (f) Reconstruction error for overall filter bank (N=32) in dB

The simulation results of the proposed method are compared with the methods of Jain–Crochiere design [9],Gradient method [26], Chen–Lee [8], Lu–Xu–Antoniou [10], Xu–Lu–Antoniou [21], [22], Sahu O.P. and Soni M.K[17], General-purpose [24], and Smith–Barnwell [15], for $N = 32$, and are summarized in Table 1. The results indicate that the performance of our proposed method is much better than all the considered methods in terms of A*s*. The proposed method also gives improved performance than General Purpose and Smith–Barnwell, methods in terms of *Ep*, than Jain–Crochiere, Gradient,Chen–Lee, Lu–Xu–Antoniou, and Xu–Lu–Antoniou methods in terms of Es, and than General-purpose and Smith–Barnwell methods in terms of linearity of the phase response.

Methods	*E*p	*E*s	*A*s (dB)	$(\in)$ (dB)	Phase response
Jain–Crochiere [9]	2.30× 10–8	1.50×10–6	33	0.015	Linear
Gradient method [26]	2.64× 10–8	3.30×10–6	33.6	0.009	Linear
General purpose [24]	0.155	6.54×10–8	49.2	0.016	Nonlinear
Smith–Barnwell [15]	0.2	1.05×10–6	39	0.019	Nonlinear
Chen–Lee [8]	2.11× 10–8	1.55×10–6	34	0.016	Linear
Lu–Xu–Antoniou [21]	1.50× 10–8	1.54×10–6	35	0.015	Linear
Xu–Lu–Antoniou [22]	3.50× 10–8	5.71×10–6	35	0.031	Linear
Sahu O.P,Soni M.K[17]	1.45× 10–8	2.76×10–6	33.913	0.0269	Linear
Proposed method	.0398	2.69×10–7	53.391	.2732	Linear

TABLE 1: Comparison of the proposed method with other existing methods based on significant parameters for N=32

According to the results obtained, some observations about filter characteristics can be made. The frequency response of H_0 for 16, 24 32 taps prototype filter shown in Figure 2. The effect of the parameter N is clearly seen on the stopband attenuation and reconstruction error of the QMF bank from the figure. Hence, longer prototype filter leads to better stopband attenuation, and better performance. As the maximum overall ripple for QMF bank decreases with increase in the length of prototype filter upto N=32. It can be noted that as the length increased to 64 there is a slight dip in the frequency response characteristic of the prototype filter which deteriorates the overall performance of the QMF bank.

5.1 APPLICATION TO SUBBAND CODING OF IMAGES

In order to assess performance of the linear phase PR QMF banks, the designed filter banks were applied for the subband coding of 256x256, and 512x512 Cameraman, Mandrill and Lena images. The criteria of comparison used is objective and subjective performance of the encoded images. Influence of certain filter characteristics, such as stopband attenuation, maximum overall ripple, and filter length on the performance of the encoded images is also investigated.

A common measure of encoded image quality is the peak signal-to-noise ratio, which is given as:

$$PSNR = 20\log_{10}(255/\sqrt{MSE}) \tag{32}$$

where *MSE* denotes the mean-squared-error

$$MSE = \frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}\left[I(x,y) - I'(x,y)\right]^2 \tag{33}$$

where, M,N are the dimensions of the image and $I(x,y)$, $I'(x,y)$ the original and reconstructed image respectively.

In general, for satisfactory reconstruction of original image, MSE must be lower, while PSNR must be high. The results of encoded Cameraman, Mandrill and Lena images are shown in figure 3. For Cameraman, Mandrill and Lena images, the best filters in sense of rate-distortion performance are QMF banks with prototype filter length greater than 16. The results obtained show that linear phase PR QMF banks are quite competitive to the best known biorthogonal filters for image coding with respect to PSNR performance for Cameraman, Mandrill and Lena images. PSNR for all three types of images increase considerably for the filter length greater than 16. The lower length affects cameraman image more as compared to other two images. Making experiments with QMF banks with the same length prototype filter and different frequency responses, filters with better stopband attenuation perform better PSNR performance.

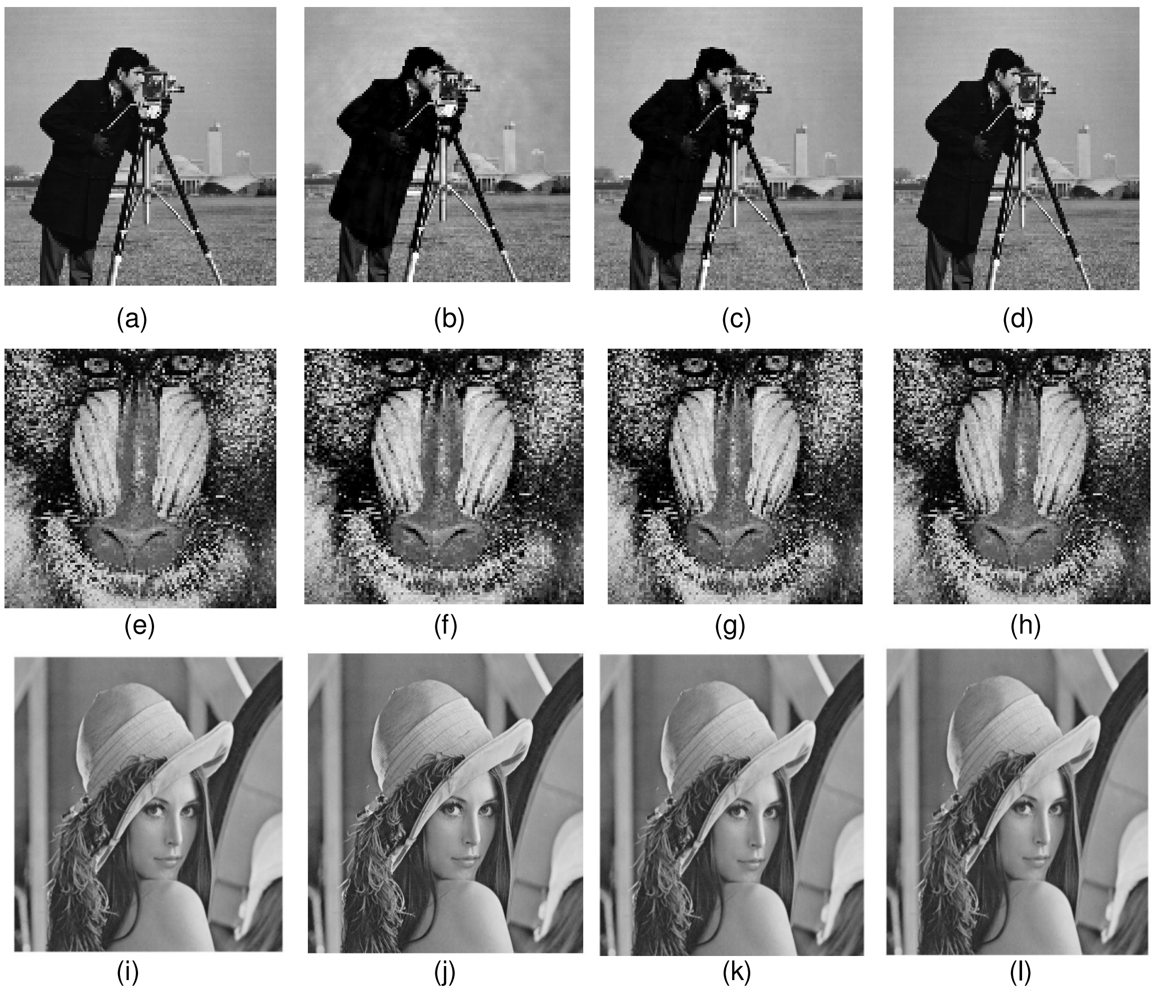


FIGURE 3: Subband coding of images using designed filter of length N (a) original cameraman image (b) reconstructed cameraman image for *N*=24 (c) reconstructed cameraman image N=32 (d) reconstructed cameraman image *N*=64 (e) original mandrill image (f) reconstructed mandrill image N=24 (g) reconstructed mandrill image N=32 (h) reconstructed mandrill image N=64 (i) original Lena image (j) reconstructed Lena image N=24 (k) reconstructed Lena image N=32 (l) reconstructed Lena image N=64

In addition to quantitative PSNR comparison, the reconstructed images were evaluated to assess the perceptual quality. For QMF banks, the perceptual quality of the image improves with the increasing length of the prototype filter h_0. This is especially obvious for the filter length above 16. The quality of encoded images obtained with QMF banks are very close to the quality of the original image. As we have been expecting, in our experiments, the most disturbing visual artifact was ringing. At lower length, this type of error affect the quality of reconstructed images significantly. The results shown in the table 2 are for single level decomposition. As the level of decomposition increases the PSNR for the Cameraman image is reduced from approximately 86 (N=24) to 75. Further, as the length increased greater than 32, complexity increased and for filter length 64 the images become brighter and the both objective and subjective performance of the images deteriorates. Thus the filter of length 24 or 32 may be used for satisfactory performance both in terms of least MSE as well as highest PSNR.

Length of prototype filter	Stop band edge	Max overall ripple dB	Stop band attenuation dB	PSNR			MSE		
				Cameraman	Mandrill	Lena	Cameraman	Mandrill	Lena
8	0.31425	8.90863	34.7831	-12	------	------	1.24e6	---	-----
12	0.309	5.80215	37.1274	52.7497	52.3128	52.8201	0.3452	0.3818	0.3397
24	0.30375	1.0089	45.8315	86.2702	83.2169	87.4151	1.5e-4	3.1e-4	1.17e-4
32	0.3025	0.30769	53.3916	89.8796	88.4642	90.2282	6.6e-5	9.26e-5	6.16e-5
64	………	0.11733	--------	53.2733	50.0073	54.3468	0.3060	0.6492	0.2390

TABLE 2: Performance results of designed QMF banks on image coding

6. CONCLUSIONS

In this paper, a modified technique has been proposed for the design of QMF bank. The proposed method optimizes the prototype filter response characteristics in passband, stopband and also the square error of the overall transfer function of the QMF bank. The method has been developed and simulated with the help of MATLAB and two design cases have been presented to illustrate the effectiveness of the proposed method. A comparison of the simulation results indicates that the proposed method gives an overall improved performance than the already existing methods, as shown in Table 1, and is very effective in designing the quadrature mirror filter banks.

We have also investigated the use of linear phase PR QMF banks for subband image coding. Coding experiments conducted on image data indicate that QMF banks are competitive with the best biorthogonal filters for image coding. The influence of certain filter characteristics on the performance of the encoded image is also analysed. It has been verified that filters with better stopband attenuation perform better rate-distortion performance. Ringing effects can be avoided by compromising between the stopband attenuation and filter length. Experimental results show that 24 and 32 taps filter is the best choice in sense of objective and subjective performances.

7. REFERENCES

1. D. Esteban, C. Galand. “Application of quadrature mirror filter to split band voice coding schemes”. *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ASSP),* pp. 191–195, 1977.

2. R.E. Crochiere. “Sub-band coding”. *Bell Syst. Tech. Journal,* 9 , pp.1633–1654,1981.

3. M.J.T. Smith, S.L. Eddins. “Analysis/synthesis techniques for sub-band image coding*”. IEEE Trans. Acoust. Speech Signal Process: ASSP*-38, pp.1446–1456, (8)1990.

4. D.Taskovski, M. Bogdanov, S.Bogdanova. “Application Of Linear-Phase Near-Perfect-Reconstruction Qmf Banks In Image Compression”. pp. 68-71, 1998.

5. J.W. Woods, S.D. O’Neil. “Sub-band coding of images”. *IEEE Trans. Acoust. Speech Signal Process.: ASSP*-34, pp.1278–1288, (10) 1986.

6. T.D. Tran, T.Q. Nguyen. "On M-channel linear phase FIR filter banks and applications in image compression". *IEEE Trans. Signal Process.*45, pp.2175–2187, (9)1997.

7. S.C. Chan, C.K.S. Pun, K.L. Ho, "New design and realization techniques for a class of perfect reconstruction two-channel FIR filter banks and wavelet bases," *IEEE Trans. Signal Process*. 52, pp.2135–2141, (7) 2004.

8. C.K. Chen, J.H. Lee. "Design of quadrature mirror filters with linear phase in the frequency domain*". IEEE Trans. Czrcuits Syst.*, vol. 39, pp. 593-605, (9)1992.

9. V.K. Jain, R.E. Crochiere. "Quadrature mirror filter design in time domain filter design in the time domain". *IEEE Trans.Acoust., Speech, Signal Processing,, ASSP*-32,pp. 353–361,(4)1984.

10. H. Xu, W.S. Lu, A. Antoniou. "An improved method for the design of FIR quadrature mirror image filter banks". *IEEE Trans. Signal Process.*46, pp.1275–1281,(6) 1998.

11. C.K. Goh, Y.C. Lim. "An efficient algorithm to design weighted minimax PR QMF banks". *IEEE Trans. Signal Process*. 47, pp. 3303–3314, 1999.

12. K. Nayebi, T.P. Barnwell III, M.J.T. Smith. "Time domain filter analysis: A new design theory". *IEEE Trans. Signal Process*. 40 ,pp.1412–1428, (6) 1992.

13. P.P. Vaidyanathan. "Multirate Systems and Filter Banks". *Prentice Hall, Englewood Cliffs, NJ*, 1993.

14. W.S. Lu, H. Xu, A. Antoniou. "A new method for the design of FIR quadrature mirror-image filter banks". *IEEE Trans. Circuits Syst. II: Analog Digital Signal Process*. 45,pp.922–927,(7) 1998.

15. M.J.T. Smith, T.P. Barnwell III. "Exact reconstruction techniques for tree structured sub-band coders*." IEEE Trans. Acoust. Speech Signal Process. ASSP*-34 ,pp.434–441, (6) 1986.

16.]J . D. Johnston. "A filter family designed for use in quadrature mirror filter banks". *Proc. IEEE Trans. Acoust, Speech, Signal Processing*, pp. 291-294, Apr. 1980.

17. O. P. Sahu, M. K. Soni and I. M. Talwar,. "Marquardt optimization method to design two-channel quadrature mirror filter banks". *Digital Signal Processing*, vol. 16,No.6, pp. 870-879, Nov. 2006.

18. H. C. Lu, and S. T. Tzeng. "Two channel perfect reconstruction linear phase FIR filter banks for subband image coding using genetic algorithm approach". *International journal of systems science*, vol.1, No. 1, pp.25-32, 2001.

19. Y. C. Lim, R. H. Yang, S. N. Koh. "The design of weighted minimax quadrature mirror filters*". IEEE Transactions Signal Processing*, vol. 41, No. 5, pp. 1780-1789, May 1993.

20. H. Xu, W. S. Lu, A. Antoniou. "A novel time domain approach for the design of Quadrature mirror filter banks". *Proceeding of the 37th Midwest symposium on circuit and system*, vol.2, No.2, pp. 1032-1035, 1995.

21. H. Xu, W. S. Lu, A. Antoniou. "A new method for the design of FIR QMF banks". *IEEE transaction on Circuits, system-II: Analog Digital processing*, vol. 45, No.7, pp. 922-927, Jul., 1998.

22. H. Xu, W.-S. Lu, A. Antoniou. "An improved method for the design of FIR quadrature mirror-image filter banks". *IEEE Transactions on signal processing*, vol. 46, pp. 1275-1281, May 1998.

23. C. W. Kok, W. C. Siu, Y. M. Law. "Peak constrained least QMF banks". *Journal of signal processing*, vol. 88, pp. 2363-2371, 2008.

24. R. Bregovic, T. Saramaki. "A general-purpose optimization approach for designing two-channel FIR filter banks". *IEEE Transactions on signal processing*, vol.51,No.7, Jul. 2003.

25. C. D. Creusere, S. K. Mitra. "A simple method for designing high quality prototype filters for M-band pseudo QMF banks". *IEEE Transactions on Signal Processing,* vol. 43, pp. 1005-1007, 1995.

26. K.Swaminathan, P.P. Vaidyanathan. “Theory and design of uniform DFT, parallel QMF banks”. *IEEE Trans. Circuits Syst.* 33 pp.1170–1191, (12) 1986.

27. A. Kumar, G. K. Singh, R. S. Anand. “Near perfect reconstruction quadrature mirror filter”. *International Journal of Computer Science and Engineering,* vol. 2, No.3, pp.121-123, Feb., 2008.

28. A. Kumar, G. K. Singh, R. S. Anand. “A closed form design method for the two channel quadrature mirror filter banks”. *Springer-SIViP,* online, Nov., 2009.

One-sample Face Recognition Using HMM Model of Fiducial Areas

OJO, John Adedapo jaojo@lautech.edu.ng
Department of Electronic & Electrical Engineering,
Ladoke Akintola University of Technology (LAUTECH),
Ogbomoso, P.M.B. 4000, Nigeria.

Adeniran, Solomon A. sadenira@oauife.edu.ng
Department of Electronic & Electrical Engineering,
Obafemi Awolowo University (OAU),
Ile Ife, Nigeria.

Abstract

In most real world applications, multiple image samples of individuals are not easy to collate for recognition or verification. Therefore, there is a need to perform these tasks even if only one training sample per person is available. This paper describes an effective algorithm for recognition and verification with one sample image per class. It uses two dimensional discrete wavelet transform (2D DWT) to extract features from images; and hidden Markov model (HMM) was used for training, recognition and classification. It was tested with a subset of the AT&T database and up to 90% correct classification (Hit) and false acceptance rate (FAR) of 0.02% was achieved.

Keywords: Hidden Markov Model (HMM); Recognition Rate (RR); False Acceptance Rate (FAR); Face Recognition (FR)

1. INTRODUCTION

Face recognition has attracted attention from the research and industrial communities with a view to achieve a "hands-free" computer controlled systems for access control, security and surveillance. Many algorithms have been proposed to solve this problem starting from the geometric feature-based [1], holistic [2,3] to appearance-based approaches [4]. The performance of these algorithms however depends heavily on the largeness of the number of training set with the attendant problem of sample collection. Algorithms that have performed excellently well with multiple sample problem (MSP) may completely fail to work if only one training sample is used [5]. However, one sample problems are more real in everyday life than the MSP. National ID cards, smart cards, student ID cards and international passports should contain enough biometric information of individuals for recognition purposes. These cases fall under the one training sample per class problem or simply one sample problem (OSP). Many algorithms have been developed and comprehensive surveys are available [5,6].

In one sample problem, the idea is to get as much information as possible from the sample. One approach to this is to increase the size of the training set by projecting the image into more than one dimension space [7], using noise model to synthesise new faces [8] or generating virtual samples or geometric views of the sample image [9]. But the one sample problem has been changed to the multiple sample problem in these cases with increase in computational and storage costs. In addition, virtual samples generated may be highly correlated and can not be considered as independent training samples [10].

In appearance-based approaches, certain features of the image samples are extracted and presented to a classifier or classifying system, which uses a similarity measure (probabilistic measure, majority voting or linearly weighted summing) to ascertain the identity of the image [11,12]. The accuracy of the performance of these methods depends largely on the features extracted from the image [5]. Gray-value features are credited with the ability to retain texture information, while Gabor and other derived features are more robust against illumination and geometrical changes [13,14]. Since there are many combining classifiers with established high level of accuracy, good performance is expected with combination of appropriate feature selection technique.

In this paper, we present a one sample face recognition and verification system, which uses two dimensional discrete Wavelets transform (2D DWT) for feature extraction and one dimensional discrete hidden Markov models (1D DHMM) for classification.

2. PRELIMINARIES

- **Hidden Markov model (HMM)**

A signal that obeys the Markov process,

$$P(q_1, q_2, \ldots q_n) = \prod_{i=1}^{T} P(q_i / q_{i-1}), \tag{1}$$

can be represented by a HMM, which consists of two interrelated processes; the observable symbol sequence and the underlying, unobservable Markov chain. A brief description of HMM is presented below, while the reader is referred to an extensive description in [15]. HMM is characterized by some elements; a specific number N of states {S}, while transition from one state S_i to another state S_j emits observation vectors O_t and the observation sequence is denoted as $O = O_1, O_2, \ldots O_T$. Observable symbols in each state can take any value in the vector $V = \{v_1, v_2, \ldots v_M\}$, where M is the number of the observable symbols. The probabilities of transition from a state i to j is expressed as,

$$a_{ij} = P\lfloor q_{t+1} = S_j / q_t = S_i \rfloor, \quad 1 \le i,\ j \le N, \tag{2}$$

$0 \le a_{ij} \le 1$ and $\sum_{j=1}^{N} a_{ij} = 1, \quad 1 \le i \le N,$

and $A = \{a_{ij}\}$

The likelihood of emitting a certain observation vector O at any state S_j is b_j, while the probability distribution $B = \{b_j(k)\}$ is expressed as,

$$b_j(k) = P\lfloor v_k \, at \; t \; / \; q_t = S_j \rfloor, \quad 1 \le j \le N, \; 1 \le K \le M \tag{3}$$

The initial state (prior) distribution $\pi = \pi_i$, where

$$\pi_i = P[q_i = S_i], \; 1 \le i \le N, \tag{4}$$

are the probabilities of S_i being the first state of sequence. Therefore a short notation for representing a model is,

$$\lambda = (A, B, \pi) \tag{5}$$

Given a model λ, and observation sequence O, the probability of the sequence given the model is $P(O \,/\, \lambda)$. This is calculated using the froward-backward algorithm,

$$P(O/\lambda) = \sum_{all\,Q} P(O/Q, \lambda) P(Q \,/\, \lambda), \tag{6}$$

$$P(O\ /\ \lambda) = \sum_{i=1}^{N} \alpha_t(i)\beta_t(i), \tag{7}$$

where $\alpha_t(i)$ is the forward variable and it is the probability of the partial observation sequence, $O = O_1, O_2, \ldots\ O_T$ and state S_j at time t, given the model λ;

$$\alpha_t(i) = P(O_1, O_2, \ldots O_t, q_t = S_i\ /\ \lambda), \tag{8}$$

$\beta_t(i)$ is the backward variable and it is is the probability of the partial observation sequence from $t+1$ to the end, given state S_i at time t and λ.

$$\beta_t(i) = P(O_{t+1}, O_{t+2}, \ldots O_T / q_t = S_i, \lambda). \tag{9}$$

The problem to solve for recognition purpose is to find the best state sequence Q that gives the maximum likelihood with respect to a given model. The viterbi algorithm is used for solving this problem. It finds the most probable path for each intermediate and finally for the terminating state. The algorithm uses two variables $\delta_t(i)$ and $\psi_t(i)$.

$$\delta_t(i) = max_{q_1, q_2, \ldots q_{t-1}}[P(q_1, q_2, \ldots\ q_t = i, O_1, O_2, \ldots\ O_t\ /\ \lambda)], \tag{10}$$

where $\delta_t(i)$ is the best score or highest probability along a single path, at time t, which accounts for the first t observations and ends in state S_j.

$$\psi_t(i) = argmax_{q_1, q_2, \ldots q_{t-1}} P(q_1, q_2, \ldots\ q_t = S_i, O_1, O_2, \ldots\ O_t\ /\ \lambda) \tag{11}$$

$\psi_t(i)$ helps to keep tract of the "best path" ending in state S_i at time t.

- **Wavelets**

Wavelet transform uses multi resolution techniques to provide a time-frequency representation of the signal. It can be described as breaking up of a signal into shifted and scaled versions of the "mother" wavelet. Wavelet analysis is done by convolving the signal with wavelet kernels to obtain wavelet coefficients representing the contributions of wavelets in the signal at different scales and orientations [16,17].

Discrete wave transform (DWT) was developed to reduce the computation time and for easy implementation of the wavelet transform. It produces a time-scaled representation of the signal by using digital filtering techniques, the wavelet families. Unlike discrete Fourier transform that can be represented by a convolution equation, DWT comprises transformation kernels or equations that differ in its expansion functions, the nature of the functions (orthogonal or bi-orthogonal basis) and how many resolutions of the functions that are computed. A signal, which passes through the filter bank shown in Figure 2 is decomposed into four lower resolution components: the approximation $(cD_{j+1}^{(a)})$, horizontal $(cD_{j+1}^{(h)})$, vertical $(cD_{j+1}^{(v)})$ and diagonal coefficients $(cD_{j+1}^{(d)})$.

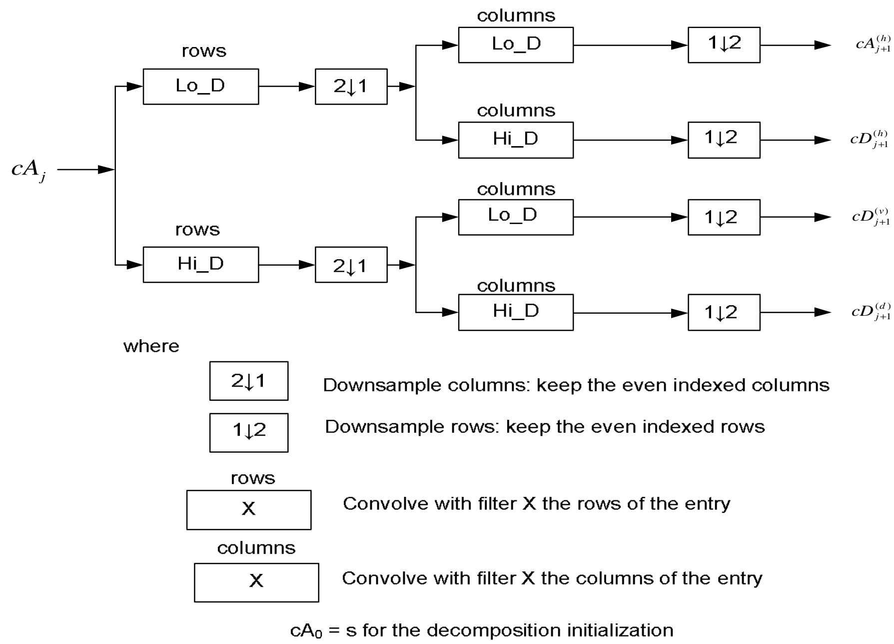


FIGURE 1: One-dimensional discrete top-to-bottom HMM (Source: MATLAB R2007a Help File)

3. METHOD

- **Modeling a Face and Feature Extraction**

A One dimensional (1D) discrete top-to-bottom (Bakis) HMM was used to segment each face into states as shown in Figure 2. The algorithm for feature extraction shown in Figure 3 was used to generate the observation vector. Two dimensional Discrete Wavelet Transform (2D DWT) was used to decompose the image into its approximation coefficients, horizontal details, vertical details and the diagonal details. ‘db1’, one of the Daubechies family of wavelets was used for decomposition. The approximation coefficient was coded using 256 gray levels thereby producing a coded (and reduced) form of the original or input image. The “coded” image was divided into sub-images and the overlap between successive sub-images was allowed to be up to 5 pixels less than the total height of the sub-image.

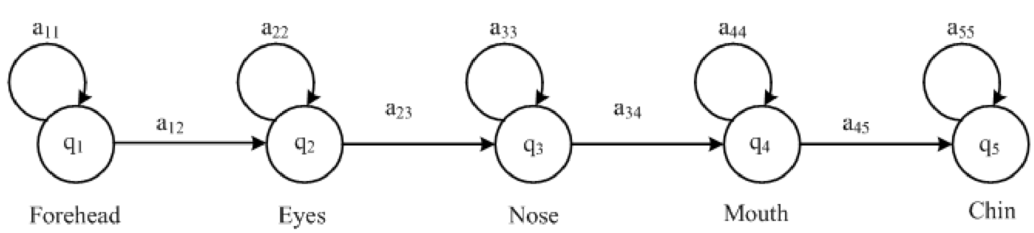


FIGURE 2: One-dimensional discrete top-to-bottom HMM

To generate the observation vector from each sub-image, the two dimensional sub-images were converted into a vector by extracting the coefficients column-wise. The number of features (NF) selected was varied to see its effect on the recognition ability of the system. The coefficients of the sub-images were stacked to form a vector, therefore a face image was represented by a vector ($Q \times NF$) in length, where Q is the number of states. Figure 4(a) shows the original image from the AT&T database while Figure 4(b) shows the gray-scale of the approximation coefficient of the same image with the sampling strip for segmenting the image into allowable states.

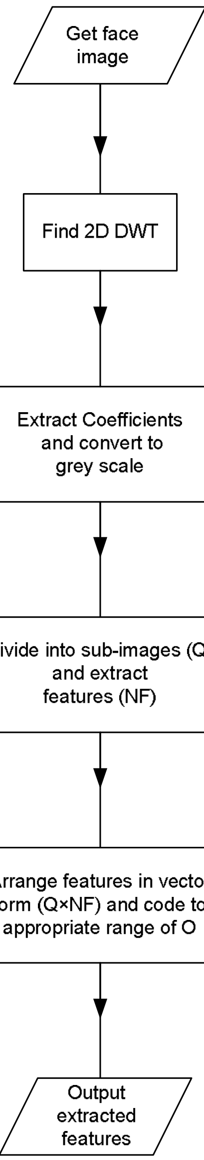


FIGURE 3: Algorithm for feature extraction

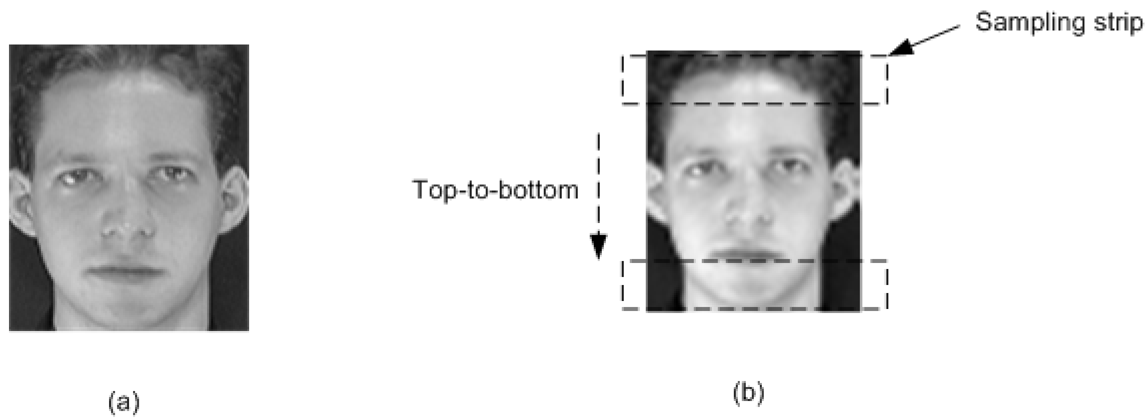


FIGURE 4: Grey-scale version of the approximation coefficients of an image and its segmentation into states

- **Training**

Features extracted from faces of individuals were used to train a model for each face using the algorithm shown in Figure 5. The initial parameter were generated randomly and improved using Baum-Welch re-estimation procedure [15] to get the parameters that optimised the likelihood of the training set observation vectors for the each face.
State transition probability (A) is defined as,

$$a_{ij} = 0, \qquad j < i \tag{12}$$
$$a_{ij} = 0, \qquad j > i + \Delta \tag{13}$$

where $\Delta = 1$ i.e. the model is not allowed to jump more than a state at a time. Since each face was divided into five sub-images, the resulting matrix is

$$A = \begin{bmatrix} a_{11} & a_{12} & 0 & 0 & 0 \\ 0 & a_{22} & a_{23} & 0 & 0 \\ 0 & 0 & a_{33} & a_{34} & 0 \\ 0 & 0 & 0 & a_{44} & a_{45} \\ 0 & 0 & 0 & 0 & a_{55} \end{bmatrix} \tag{14}$$

$a_{NN} = p$, while $a_{Ni} = 0 \;\; for\; i < N\; and\; i > N+1$ and the initial state probability (π) is defined as

$$\pi_i = \begin{cases} 0, & j \neq 1 \\ 1, & j = 1 \end{cases} \tag{15}$$

$$\pi_i = [1,0,0,0,0,0]. \tag{16}$$

The maximum number of iteration for the re-estimation is set to 5 or if the error between the initial and present value is less than 10^{-4}, then the model is taken to have converged and the model parameters are stored with appropriate class name or number (A_c, B_c, π_c).

- **Algorithm for Model Re-estimation**

(n is the maximum number of iteration allowed)
k = 1
initialise $\lambda = (A, B, \pi)$
compute $P(O/\lambda^k)$
while $k < n$ do
    estimate $P(O/\lambda^{k+1})$
    if $|P(O/\lambda^{k+1}) - P(O/\lambda^k)| < error$
        quit
    else
        $P(O/\lambda^k) \Leftarrow P(O/\lambda^{k+1})$
    end
end

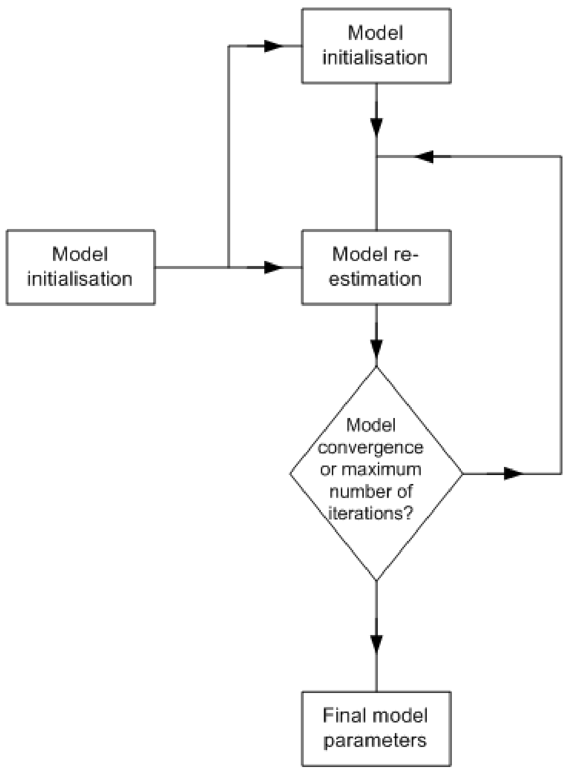


FIGURE 5: Algorithm for HMM training

- **Recognition and Verification**

Given a face to be tested or recognised, feature (observation vector) extraction is first performed as described in section 3.1. Model likelihoods (log-likelihood) for all the models in the training set (given the observation vectors) is calculated and the model with the highest log-likelihood is identified to be the model representing the face. Euclidean measure is used to test if a face is in the training set or database. If the log-likelihood is within the stated distance, the model (face) is recognised to be in the training set or in the database. However, in areas of applications such as access control, it is desired to know the exact identity of an individual, therefore the need to verify the correctness of the faces recognised.

For classification or verification, the Viterbi recogniser was used as shown in Figure 6. The test (face) image was converted to an observation sequence and then model likelihoods $P(O_{test} / \lambda_i)$

are computed for each $\lambda_i, i = 1,2, \dots c$. The model with highest likelihood reveals the identity of the unknown face.

$$v = argmax_{1 \le i \le c}[P(O_{test} / \lambda_i)] \quad (12)$$

4. RESULTS AND DISCUSSION

The algorithm was implemented in Matlab 7.4a on a HP AMD Turion 64 X2 TL-58, 1.90GHz, 2GB RAM on a Windows operating system. It was tested with a subset of the AT&T (formerly ORL) database [18]. A face image per person was used for training while five other images per person were used for testing, some of which are shown in Figure 7. The recognized images were verified for correctness, 80% correct classification (Hit) occurred while 20% were misclassified. The rest of the images that were not in the training set were used to test the false acceptance rate (FAR) i.e. the ratio of the numbers of images falsely accepted to the total number of images tested and 0.02 FAR occurred. The number of test images per class was reduced to two and 90% Hit, 0.025 FAR occurred as shown in Table 1. Furthermore, the algorithm was tested with ten subjects in the AT&T database. The general observation was that the percentage Hit and FAR were independent of number of subjects in class. For instance, 90% Hit, 0.05 FAR and 90% Hit, 0.05 FAR occurred when five and two test images per class were used respectively.

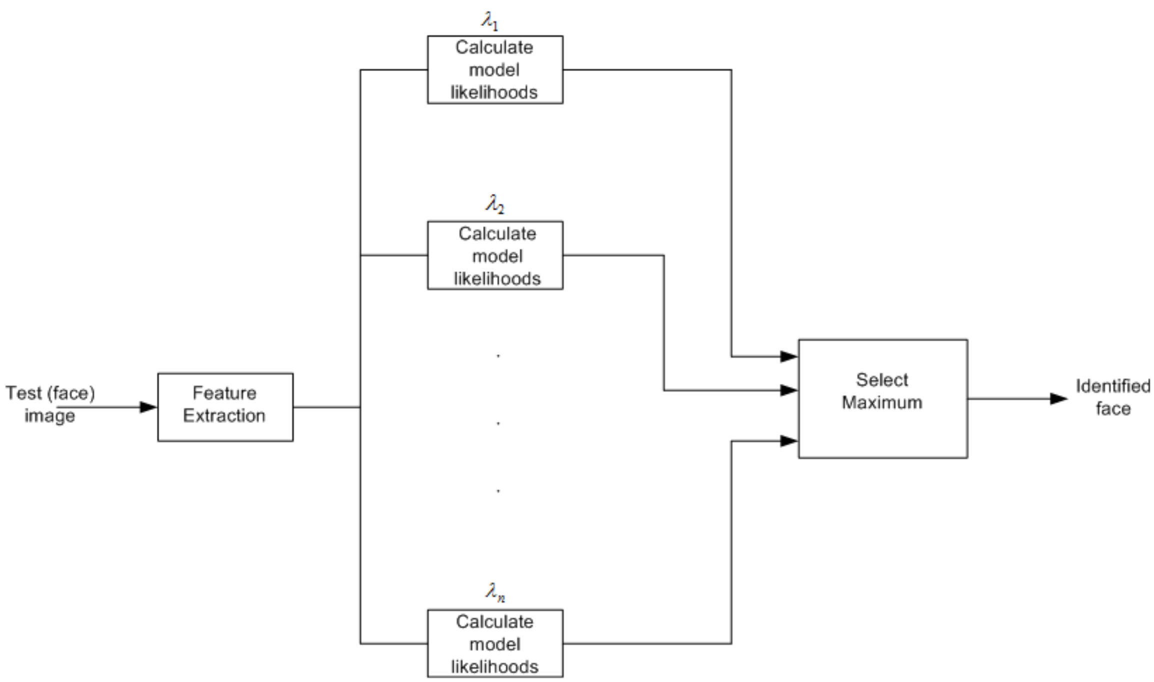


FIGURE 6: Viterbi recogniser

Figure 8 shows the effect that the number of features selected per each state (subimage) has on the number of correct classifications (Hit). The results show that 30 features per subimage were sufficient to give the best performance. In addition, the average time for testing a face was approximately 0.15s, which is near real-time. Going by these results, the algorithm is expected to be adequate for implementation in applications where small size database is required [19].

The performance of the algorithm when Two Dimensional Discrete Cosine Transform (2D-DCT) was used for feature extraction is compared with that of the Discrete Wavelet Transform (2D-DWT) and the results are shown in Table 2. The results show that there is a significant improvement in the recognition or classification accuracy when DWT was used for feature extraction.

Test Images	Number of class	Hit	Miss	FAR
5	20	80%	20%	0.02
2	20	90%	10%	0.025
5	10	90%	10%	0.04
2	10	90%	10%	0.05

TABLE 1: Classification accuracies achieved for a subset of AT&T database

Test Images	Number of class	Hit	
		DCT	DWT
5	20	39%	80%
2	20	50%	90%
5	10	46%	90%
2	10	45%	90%

TABLE 2: Classification accuracies for DCT and DWT

FIGURE 7: Some of the faces used for testing.

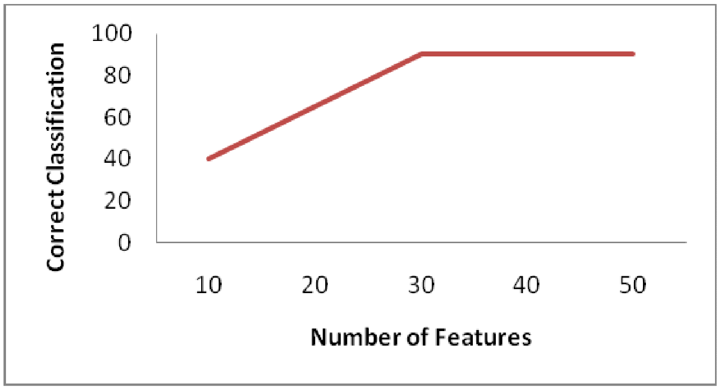


FIGURE 8: Graph of correct classifications against number of features per state

5. CONSLUSION

The paper presented a one sample face recognition system. Feature extraction was performed using 2D DWT and 1D top-to-bottom HMM was used for classification. When tested with a subset of the AT&T database, up to 90% correct classification (Hit) and as low as 0.02 FAR were achieved. The high recognition rate and the low FAR achieved shows that the new algorithm is suitable for face recognition problems with small-size database such as access control for personal computers (PCs) and personal digital assistants (PDAs).

6. REFERENCES

1. R. Brunelli and T. Poggio. *"Face recognition: Features versus templates"*. IEEE Transaction on Pattern Analysis and Machine Intelligence, 15(10):1042-1062, 1993

2. L. Sirovich and M. Kirby. *"Low-Dimensional procedure for the characterization of human face"*. Journal of the Optical Society of America A, 4(3):519–524, 1987

3. M. Turk and A. Pentland. *"Eigenfaces for Recognition".* Journal of Cognitive Neuroscience, 3(1):71-86, 1991

4. S. Lawrence, C.L. Giles, A. Tsoi and A. Back. *"Face recognition: A convolutional neural-network approach"*. IEEE Transaction on Neural Networks, 8(1):98-113, 1997

5. W. Zhao, R. Chellappa, P.J. Philips and A. Rosenfeld. *"Face recognition: A literature survey".* ACM Computing Surveys 35(4):399-458, 2003

6. X. Tan, S. Chen, Z-H Zhou, and F. Zhang. *"Face recognition from a single image per person: a survey".* Pattern Recognition 39:1725-1745, 2006

7. J. Wu and Z.-H Zhou. *"Face recognition with one training image per person".* Pattern Recognition Letters, 23(2):1711-1719, 2001

8. H.C. Jung, B.W. Hwang and S.W. Lee. *"Authenticating corrupted face image based on noise model".* Proceedings of the 6th IEEE International Conference on Automatic Face and Gesture Recognition, 272, 2004

9. F. Frade, De la Torre, R. Gross, S. Baker, and V. Kumar. *"Representational oriented component analysis (ROCA) for face recognition with one sample image per training class".* Proceedings of IEEE Conference on Computer Vision and Pattern Recognition 2:266-273, June 2005

10. A.M. Martinez. *"Recognizing imprecisely localised, partially occluded, and expression variant faces from a single sample per class".* IEEE Transaction on Pattern Analysis and Machine Intelligence 25(6):748-763, 2002

11. B.S. Manjunath, R. Chellappa and C.V.D. Malsburg. *"A feature based approach to face recognition".* In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition 1:373-378, 1992

12. M. Lades, J.Vorbruggen, J. Buhmann, J. Lange, von der Malsburg and R. Wurtz. *"Distortion invariant object recognition in the dynamic link architecture".* IEEE Transaction on Computers 42(3):300-311, 1993

13. X. Tan, S.C. Chen, Z-H Zhou, and F. Zhang. *"Recognising partially occluded, expression variant faces from single training image per person with SOM and soft kNN ensemble".* IEEE Transactions on Neural Networks, 16(4):875-886, 2005

14. H.-S. Le and H. Li. *"Recognising frontal face images using Hidden Markov models with one training image per person".* Proceedings of the 17th International Conference on Pattern Recognition (ICPR04), 1:318-321, 2004

15. L.R. Rabiner. *"A tutorial on Hidden Markov models and selected application in speech recognition".* Proceedings of the IEEE, 77(2):257-286, 1989

16. I. Daubechies, *"Orthonormal bases of compactly supported wavelets".* Communication on Pure & Applied Mathematics XLI, 41:909-996, 1988

17. I. Daubechies. *"Ten lectures on wavelets".* CBMS-NSF Conference series in Applied Mathematics, No-61, SIAM, Philadelphia Pennsylvania, 1992

18. F. Samaria and A. Harter. *"Parameterization of a stochastic model for human face identification".* 2nd IEEE Workshop on Applications of Computer Vision, Saratosa FL. pp.138-142, December, 1994

19. J. Roure, and M. Faundez-Zanuy. *"Face recognition with small and large size databases".* Proceedings of 39th Annual International Carnahan Conference on Security Technology, pp 153-156, October 2005

A New Method Based on MDA to Enhance the Face Recognition Performance

Aref Shams Baboli Aref.shams2222@gmail.com
BSc *Student at Department of Electrical and computer Engineering, Noshirvani University of Technology, mazandaran, Iran*

Seyyedeh Maryam Hosseyni Nia maryamhosseini16@yahoo.com
MSc Student at Department of Biomedical Engineering and Medical Physics Shahid Beheshti University, Tehran, Iran

Ali Akbar Shams Baboli ali.shams2222@gmail.com
MSc Student at Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

Gholamali Rezai Rad rezai@iust.ac.ir
Associate professor at Department of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

Abstract

A novel tensor based method is prepared to solve the supervised dimensionality reduction problem. In this paper a multilinear principal component analysis (MPCA) is utilized to reduce the tensor object dimension then a multilinear discriminant analysis (MDA), is applied to find the best subspaces. Because the number of possible subspace dimensions for any kind of tensor objects is extremely high, so testing all of them for finding the best one is not feasible. So this paper also presented a method to solve that problem, the main criterion of algorithm is not similar to Sequential mode truncation (SMT) and full projection is used to initialize the iterative solution and find the best dimension for MDA. This paper is saving the extra times that we should spend to find the best dimension. So the execution time will be decreasing so much. It should be noted that both of the algorithms work with tensor objects with the same order so the structure of the objects has been never broken. Therefore the performance of this method is getting better. The advantage of these algorithms is avoiding the curse of dimensionality and having a better performance in the cases with small sample sizes. Finally, some experiments on ORL and CMPU-PIE databases are provided.

Keywords: Dimensionality Reduction, *HOSVD*, Subspace Learning, Multilinear Principal Component Analysis, Multilinear Discriminant Analysis.

1. INTRODUCTION

A typical tensor object in machine vision or pattern recognition applications is actually in a high-dimensional tensor space. In reality, the extracted features of an object often has some specific structures that are in the form of second or even higher order tensors [1]. Most previous works transform the input image data into a 1-D vector, which ignores the underlying data structure so these methods suffer from *curse of dimensionality* and *small sample size problem.* Subspace learning is one of the most important directions in computer vision research [2], [3]. Most traditional algorithms, such as LDA [4] input an image object as a 1-D vector. It is well understood that reshaping breaks the natural structure and correlation in the original data.

Some recent works have started to consider an object as a 2-D matrix rather than vectors for subspace learning. A 2-D PCA algorithm is proposed in [5] where gets the input images as a matrix and compute a covariance matrix. As we mentioned before, in this paper a method that utilized the MDA after MPCA algorithms has been proposed in which both of those algorithms work with tensor objects that give us the better results.
It should be noted that recently there are many developments in the analysis of higher order. Reference [6] used a MPCA method based on HOSVD [7]. There is also a recent work on multilinear discriminant analysis (MDA) in [8] which is used for maximizing a tensor-based discriminant criterion. Previously, we proposed MPCA+MDA [9] for face recognition. In that paper we use MPCA algorithm for tensor object feature extraction. MPCA is a multilinear algorithm reducing dimension in all tensor modes to find those bases in each mode that allows projected tensors to achieve most of the original tensors variation. Then these bases are applied on samples and a new data set with a new dimension will be generated. This new data set will be the inputs of our MDA algorithm. MDA uses a novel criterion for dimensionality reduction, *discriminant tensor criterion* (DTC), which maximizes the interclass scatter and simultaneously time minimizes the intraclass scatter. In that paper we should give the goal dimension for reduction manually. As we know, the number of possible subspace dimensions for tensor objects is extremely high, so testing all of them to find the best one is not feasible. To solve that problem, a method is used to find the best dimension that gives us the best accuracy. Our method is a little similar to SMT that is used in MPCA algorithm [5]. To start the algorithm like SMT we need to initialization the subspaces. So this paper is used full projection to initialize the iterative solution for MDA [6]. The main idea of this paper is saving the extra times that we should spend to find the best dimension and of course the final dimension in MDA that we find practically is not optimal. But with our improvement we are decreasing the execution time so much.
MPCA+Improved MDA can avoid the *curse of dimensionality dilemma* by using higher order tensor for objects and *n-mode optimization* approach. Due to using the MDA after applying the MPCA, this method is performed in a much lower-dimension feature space than MDA and the traditional vector-based methods, such as LDA and PCA do. Also because of the structure of MDA, it can overcome the *small sample size* problem. As we know, the available feature dimension of LDA is theoretically limited by the number of classes in the data set but in our algorithm it is not limited. So it would give us the better recognition accuracy. As a result of all the above characteristics, we expect this novel method to be a better choice than LDA and PCA algorithms and more general than MDA for the pattern classification problems in image analysis and also overcome the small sample sizes and curse of dimensionality dilemma.
The rest of this paper is organized as follows. Section 2 introduces basic multilinear algebra notations and concepts. In Section 3, the Initialization procedures of MPCA and introducing the DTC and *n-mode optimization* that are used in MDA is discussed after that we will see our proposed method for finding the best subspaces dimension. Then, in Section 4, we present the face recognition experiments by encoding the image objects as second or third-order tensors and compare them to traditional subspace learning algorithms and MDA algorithm. Finally, in Section 5, the major point of this paper and the future work is summarized.

2. MULTILINEAR NOTATIONS AND BASIC ALGEBRA

This section briefly will be reviewed some basic multilinear concepts used in our framework and see an example for *n–mode unfolding* of a tensor. Here, vectors are denoted by lowercase boldface letters, such as, **x**, **y**. The bold uppercase symbols are used for representing matrices, such as **U**, **S**, and tensors by calligraphic letters, e.g. $\mathcal{A}$. An Nth-order tensor is denoted as $\mathrm{A} \in \mathrm{R}^{I_1 \times I_2 \times \ldots \times I_N}$. It is addressed by N indices I_n, n = 1, … , N and each I_n addresses the *n*-mode of $\mathcal{A}$. The *n*-mode product of a tensor $\mathcal{A}$ by a matrix $\mathbf{U}$, is

$$\left(\mathrm{A} \times_n \mathbf{U}\right)\left(i_1,\ldots,i_{n-1}, j_n, i_{n+1},\ldots,i_N\right) = \sum_{i_n} \mathrm{A}\left(i_1,\ldots,i_N\right).\mathbf{U}(j_n, i_n) \tag{1}$$

The scalar product of two tensors $\mathcal{A}, \mathcal{B} \in R^{I_1 \times I_2 \times .. \times I_N}$ is defined as $\prec \mathcal{A}, \mathcal{B} \succ = \sum_{i_1}\sum_{i_2}\cdots\sum_{i_N} \mathcal{A}\left(i_1,\ldots,i_N\right).\mathcal{B}(i_1,\ldots,i_N)$ and the Frobenius norm of $\mathcal{B}$ is defined as $\mathcal{B}_F = \sqrt{\prec \mathcal{B}, \mathcal{B} \succ}$ [7].

Unfolding along the *n*-mode is denoted as $\mathbf{A}_{(n)} \in R^{I_n \times (I_1 \times .. \times i_{n-1} \times i_{n+1} \times .. \times i_N)}$. The column vectors of $\mathbf{A}_{(n)}$ are the *n*-mode vectors of $\mathcal{A}$. Fig. 1 illustrates three ways to unfold a third-order tensor. For unfolding along the first-mode, a tensor is unfolded into a matrix along the I_1 axis, and the matrix width direction is indexed by searching index I_2 and I_3 index iteratively. In the second-mode, the tensor is unfolded along the I_2 axis and the same trend afterwards.

Following standard multilinear algebra, tensor $\mathcal{A}$ can be expressed as the product $\mathcal{A} = \mathcal{S} \times_1 \mathbf{U}^{(1)} \times_2 \mathbf{U}^{(2)} \times \ldots \times_N \mathbf{U}^{(N)}$. Where $\mathcal{S} = \mathcal{A} \times_1 \mathbf{U}^{(1)\mathrm{T}} \times_2 \mathbf{U}^{(2)\mathrm{T}} \times \ldots \times_N \mathbf{U}^{(N)\mathrm{T}}$ and we call $\mathcal{S}$ core tensor that will be used for HOSVD and $\mathbf{U}^{(n)} = \left(\mathbf{u}_1^{(n)}\mathbf{u}_2^{(n)} \ldots \mathbf{u}_{I_n}^{(n)}\right)$ is an orthogonal $I_n \times I_n$ matrix. The relationship between unfolded tensor $\mathbf{A}_{(n)}$ and its decomposition core tensor $\mathcal{S}_{(n)}$ is

$$\begin{aligned}\mathbf{A}_{(n)} &= \mathbf{U}^{(n)}.\mathcal{S}_{(n)}.(\mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \otimes \ldots \\ &\otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \ldots \otimes \mathbf{U}^{(n-1)})^{\mathrm{T}}\end{aligned} \tag{2}$$

Where $\otimes$ means the Kronecker product [7].

The projection of an n-mode vector of $\mathcal{A}$ by $\mathbf{U}^{(n)\mathrm{T}}$ is computed as the inner product between the n-mode vector and the rows of $\mathbf{U}^{(n)\mathrm{T}}$. For example in Fig. 2, a third-order tensor $\mathcal{A} \in R^{I_1 \times I_2 \times I_3}$ is projected in the 1-mode vector space by a projection matrix $\mathbf{B}^{(1)\mathrm{T}} \in R^{m_1 \times I_1}$, the projected tensor is $\mathcal{A} \times_1 \mathbf{B}^{(1)\mathrm{T}} \in R^{m_1 \times I_2 \times I_3}$. In the 1- mode projection, each 1-mode vector of length I_1 is projected by $\mathbf{B}^{(1)\mathrm{T}}$ to obtain a vector of length m_1 .

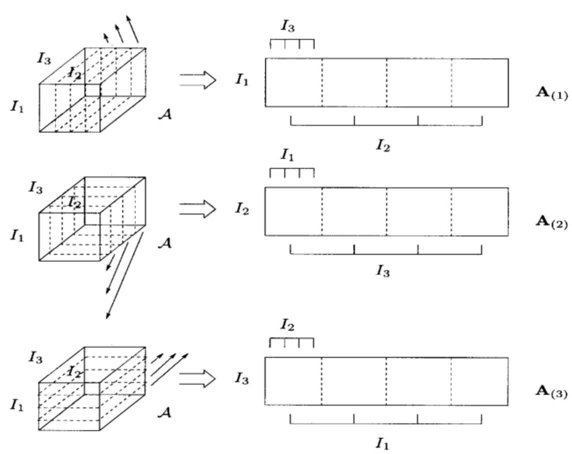


FIGURE 1: Illustration of the *n-mode unfolding* of a third–order tensor.

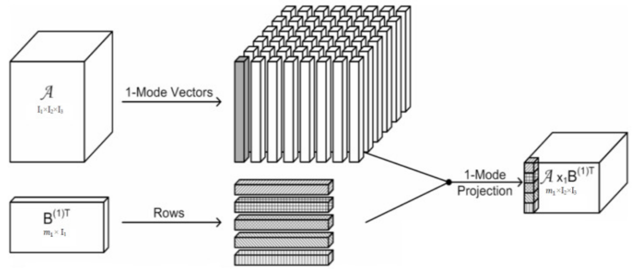


FIGURE 2: Illustration of multilinear projection in the mode 1

3. MULTILINEAR PRINCIPAL COMPONENT ANALYSIS & MULTILINEAR DISCRIMINANT ANALYSIS

Some previous approaches to subspace learning, such as PCA and LDA, consider an object as a 1-D vector so the learning algorithms should be applied on a very high dimension feature space. So these methods suffer from the problem of *curse of dimensionality*. Most of the objects in computer vision are more naturally represented as second or higher order tensors. For example, the image matrix in Fig. 3(a) is a second-order tensor and the filtered Gabor image in Fig. 3(b) is a third-order tensor.

In this section, first we see, how the MPCA solution for tensor objects is working and then we will see the DTC and *n-mode optimization* that is used in MDA for tensor objects. A set of M tensor objects $\{\mathrm{X}_1, \mathrm{X}_2, ..., \mathrm{X}_M\}$ is available for training. Each tensor object $\mathrm{X}_m \in \mathrm{R}^{I_1 \times I_2 \times ... \times I_N}$ assumes values in a tensor space $\mathrm{R}^{I_1} \otimes \mathrm{R}^{I_2} \ldots \otimes \mathrm{R}^{I_N}$, where I_n is the *n*-mode dimension of the tensor. The MPCA defines a multilinear transformation that maps the original tensor space into a tensor subspace. In other words, the MPCA objective is the determination of the projection matrices $\left\{\mathbf{U}^{(n)} \in \mathrm{R}^{I_n \times P_n}, n = 1, \ldots, N\right\}$ that maximize the total tensor scatter, Ψ_{y}

$$\{\mathbf{U}^{(n)}, n = 1, \ldots, N\} = \arg\arg \max_{\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(N)}} \max_{\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(N)}} \Psi_{\mathrm{y}} \tag{3}$$

Where $\Psi_{\mathrm{y}} = \sum_{m=1}^{M} \left\| \mathrm{A}_m - \overline{\mathrm{A}} \right\|_F^2, \overline{\mathrm{A}} = (1/m) \sum_{m=1}^{M} \mathrm{A}_m$.

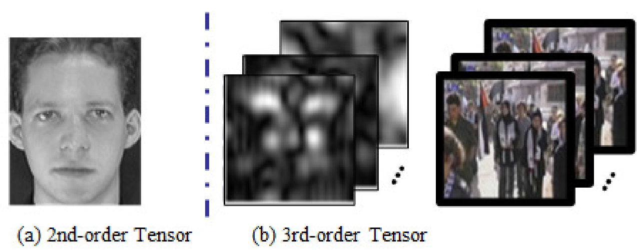


FIGURE 3: Second- and third-order Tensor representations samples

3.1 MPCA Algorithm

There is no optimal solution for optimizing the *N* projection matrices simultaneously. An *N*th-order tensor consists of N projections with *N* matrix, so *N* optimization subproblems can be solved by finding the $\mathbf{U}^{(n)}$ that maximizes the scatter in the *n*-mode vector subspace. If $\left\{\mathbf{U}^{(n)}, n = 1, \ldots, N\right\}$ be the answer of (3) and $\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(n-1)}, \mathbf{U}^{(n+1)}, \ldots, \mathbf{U}^{(N)}$ be all the other known projection matrices, the matrix $\mathbf{U}^{(n)}$ consists of the P_n eigenvectors corresponding to the largest eigenvalues of the matrix $\Phi^{(n)}$

$$\Phi^{(n)} = \sum_{m=1}^{M} \left(\mathbf{X}_{m(n)} - \overline{\mathbf{X}}_{(n)}\right).\mathbf{U}_{\Phi^{(n)}}.\mathbf{U}_{\Phi^{(n)}}^{T}.\left(\mathbf{X}_{m(n)} - \overline{\mathbf{X}}_{(n)}\right)^{T} \tag{4}$$

Where $\mathbf{U}_{\Phi^{(n)}} = \mathbf{U}^{(n+1)} \otimes \mathbf{U}^{(n+2)} \otimes \ldots \otimes \mathbf{U}^{(N)} \otimes \mathbf{U}^{(1)} \otimes \mathbf{U}^{(2)} \otimes \ldots \mathbf{U}^{(n-1)}$.
The proof of (4) is given in [6].
Since $\Phi^{(n)}$ depends on all the other projection matrices, there is no closed-form solution for this maximization problem. Instead, reference [6] introduce an iterative procedure that can be utilized to solve (4). For initialization, MPCA used full projection. The term full projection refers to the multilinear projection for MPCA with $P_n=I_n$ for n= 1, …, N. There is no dimensionality reduction through this full projection. The optimal is obtained without any iteration, and the total scatter in the original data is fully captured. After finding the projection matrices, $\mathbf{U}^{(n)}, n = 1,\ldots,N,$ we applied those matrices to the training set. At this point, we provide a set of tensors with the new dimension that would be the new training set for MDA algorithm.

3.2 Multilinear Discriminant Analysis

Here, the DTC is introduced which is used in MDA algorithm. The DTC is designed to provide multiple interrelated projection matrices, which maximize the interclass scatter and at the same time minimize the intraclass scatter. That is

$$\mathbf{U}^{(n)*}\Big|_{n=1}^{N} = \arg\max_{\mathbf{U}^{(n)}|_{n=1}^{N}} \frac{\sum_c n_c \overline{\mathrm{X}}_c \times_1 \mathbf{U}^{(1)} \ldots \times_n \mathbf{U}^{(n)} - \overline{\mathrm{X}} \times_1 \mathbf{U}^{(1)} \ldots \times_n \mathbf{U}^{(n)2}}{\sum_i \mathrm{X}_i \times_1 \mathbf{U}^{(1)} \ldots \times_n \mathbf{U}^{(n)} - \overline{\mathrm{X}}_{c_i} \times_1 \mathbf{U}^{(1)} \ldots \times_n \mathbf{U}^{(n)2}}. \tag{5}$$

Where $\overline{\mathrm{X}}_c$ is the average tensor of class c samples, $\overline{\mathrm{X}}$ is the total average tensor of all the samples, and n_c is sample number of class *c.* We could optimize that function by using *n-mode optimization* approach that is proved in [8]. The optimization problem can be reformulated as follows:

$$\mathbf{U}^{(n)*} = \arg\max_{\mathbf{U}^{(n)}} \frac{Tr(\mathbf{U}^{(n)T} S_B \mathbf{U}^{(n)})}{Tr(\mathbf{U}^{(n)T} S_W \mathbf{U}^{(n)})} \tag{6}$$

$$S_B = \sum_{j=1}^{\prod_{o\neq n} m_o} S_B^j, S_B^j = \sum_{c=1}^{N_c} n_c (\overline{\mathbf{X}}_{c(n)}^{\,j} - \overline{\mathbf{X}}_{(n)}^{\,j})(\overline{\mathbf{X}}_{c(n)}^{\,j} - \overline{\mathbf{X}}_{(n)}^{\,j})^T$$

$$S_W = \sum_{j=1}^{\prod_{o\neq n} m_o} S_W^j, S_W^j = \sum_{i=1}^{N_c} (\overline{\mathbf{X}}_{i(n)}^{\,j} - \overline{\mathbf{X}}_{c_i(n)}^{\,j})(\overline{\mathbf{X}}_{i(n)}^{\,j} - \overline{\mathbf{X}}_{c_i(n)}^{\,j})^T$$

Where, $\mathbf{X}_{i(n)}^{\,j}$ is the *j*th column vector of matrix $\mathbf{X}_{i(n)}^{\,j}$ which is the *n*-mode unfolded matrix from sample tensor X_i . $\mathbf{X}_{c(n)}^{\,j}$ and $\overline{\mathbf{X}}_{(n)}^{\,j}$ are defined in the same way as $\mathbf{X}_{i(n)}$ with respect to tensors $\overline{\mathrm{X}}_c$ and $\overline{\mathrm{X}}$ and the proofs are given in [8]. To utilizing *n-mode optimization*, first the input tensors (that are the outputs of MPCA) should be projected with all the other modes matrices and then all the new tensors are unfolded into a matrix along the nth-mode. Therefore, the optimization problem in (5) can be reformulated as a special discriminant analysis problem, and it can be solved in the same way for the traditional LDA algorithm [8]. Since DTC has no closed form the projection matrices can be iteratively optimized.

3.3 Determination of the Tensor Subspace Dimensionality

The target dimensionality P_n has to be determined. So the objective MPCA function should be revised to include a constraint on the favorite dimensionality reduction. The revised objective function is as follows [6]:

$$\left\{\mathbf{U}^{(n)}, P_n, n=1,\ldots,N\right\}$$

$$= \arg \max_{\mathbf{U}^{(1)},\mathbf{U}^{(2)},\ldots,\mathbf{U}^{(N)},P_1,P_2,\ldots,P_N} \frac{Tr(\mathbf{U}^{(n)T} S_B \mathbf{U}^{(n)})}{Tr(\mathbf{U}^{(n)T} S_W \mathbf{U}^{(n)})}$$

$$\text{subject to} \frac{\prod_{n=1}^{N} P_n}{\prod_{n=1}^{N} I_n} < \Omega \tag{7}$$

Where the ratio between the reduced dimensionality and the original tensor space dimensionality is utilized to measure the amount of dimensionality reduction, and Ω is a threshold to be specified by user.

The proposed tensor subspace dimensionality determination solution is Starting with $P_n=I_n$ for all n at $t=0$, at each subsequent step $t=t+1$, this algorithm truncates, in a selected mode n, the P_n th n-mode eigenvector of the reconstructed input tensors. The truncation can be interpreted as the elimination of the corresponding P_n th n-mode slice of the total scatter tensor. For the specific mode selection, the scatter loss rate $\delta_t^{(n)}$ due to the truncation of its P_n th eigenvector is calculated for each mode. $\delta_t^{(n)}$ is defined as follows [6]:

$$\delta_t^{(n)} = \frac{\dfrac{Tr(\mathbf{U}^{(n)T} S_B \mathbf{U}^{(n)})}{Tr(\mathbf{U}^{(n)T} S_W \mathbf{U}^{(n)})}_{\mathrm{Y(t)}} - \dfrac{Tr(\mathbf{U}^{(n)T} S_B \mathbf{U}^{(n)})}{Tr(\mathbf{U}^{(n)T} S_W \mathbf{U}^{(n)})}_{\mathrm{Y(t-1)}}}{\left[P_n . \prod_{j=1, j\neq n}^{N} P_j\right] - \left[(P_n - 1) . \prod_{j=1, j\neq n}^{N} P_j\right]}$$

$$= \frac{\lambda_{P_n}^{(n)}}{\prod_{j=1, j\neq n}^{N} P_j} \tag{8}$$

Where $Tr(\mathbf{U}^{(n)T} S_B \mathbf{U}^{(n)}) \big/ Tr(\mathbf{U}^{(n)T} S_W \mathbf{U}^{(n)})_{\mathrm{Y(t)}}$ is maximizing the between class scatter and at the same time minimizing the within class scatter at step t, $\prod_{j=1, j\neq n}^{N} P_j$ is the amount of dimensionality reduction achieved, and $\lambda_{P_n}^{(n)}$, which is the corresponding P_n th n-mode eigenvalue, is the loss due to truncating the P_n th n-mode eigenvector. The mode with the smallest $\delta_t^{(n)}$ is selected for the step-t truncation. For the selected mode n, P_n is decreased by 1: $P_n = P_n - 1$ and $\prod_{n=1}^{N} P_n \big/ \prod_{n=1}^{N} I_n$ is tested. The truncation stops when $\prod_{n=1}^{N} P_n \big/ \prod_{n=1}^{N} I_n$ is satisfied. The term full projection refers to the multilinear projection for MDA with $P_n=I_n$ for n= 1, … , N for starting the algorithm. There is no dimensionality reduction through this full projection [5]. The optimal is obtained without any iteration. As we know, if all eigenvalues (per mode) are distinct, the full projection matrices are also distinct. Therefore, the full projection is unique [6].

4. EXPERIMENTS

In this section, two standard face databases ORL [10], CMU PIE [11] were used to evaluate the effectiveness of our proposed algorithm, MPCA+Improved MDA, in face recognition accuracy. These algorithms were compared with the popular Eigenface, Fisherface and MDA/2-1, MDA/2-2, MDA/3-3 and the MPCA + MDA algorithms. In this work, we report the best result on different test and for the fisherface on different feature dimensions in the LDA step, in all the experiments, the training and test data were both transformed into lower dimensional tensors or vectors via the learned subspaces, and we use the nearest neighbour classifier for final classification. The performances on the cases with different number of training samples were also evaluated to illustrate their robustness in the small sample size problems.

4.1 ORL Database

The ORL database includes 400 images of 40 persons. These images were captured at different times and have different expression such as open or closed eyes, smiling or nonsmiling and facial details like: glasses or no glasses. All images were in grayscale and centered with the resolution of 112*92 pixels. Ten sample images of one person in the ORL database are displayed in Figure

FIGURE 4: Ten samples of one person in the ORL face database

Four sets of experiments were managed to compare the performance of our algorithm with Eigenface, Fisherface, and MDA/2-1, MDA/2-2. In each experiment, the image set was partitioned into the test and train set with different numbers. Table 1 shows the best face recognition accuracies of all the algorithms in our experiments with different train and test set partitions. The results show that our algorithm outperforms Eigenface, Fisherface, MDA/2-1, MDA/2-2 and MPCA+MDA on all four sets of experiments, especially in the cases with a small number of training samples and also we can see the performance of MPCA+Improved MDA is the same as MPCA+MDA or even better than that. It means we provide the same performance without spending the spare time to find the best dimension. So we can say our very new proposed algorithm has the best performance and also save the spare times.

TABLE 1: Recognition Accuracy (%) Comparsion of MDA+MPCA, Eigenface, Fisherface, MDA/2-1, MDA/2-2, MPCA+MDA on ORL database

Algorithms	Train-Test			
	5-5	***4-6***	***3-7***	***2-8***
Eigenface	97.0	91.25	87.50	81.56
Fisherface	93.0	85.83	87.50	79.68
MDA/2-1	97.5	96.25	94.28	88.13
MDA/2-2	99.0	97.91	95.00	90.31
MPCA + MDA	99.0	98.75	96.43	91.56
MPCA + Improved MDA	99.0	98.75	96.78	91.87

4.2 CMU PIE Database

The CMU PIE database contains more than 40,000 facial images of 68 people. The images were obtained over different poses, under variable illumination conditions and with different facial expressions. In our experiment, two sub-databases were used to evaluate our methods. In the first sub-database, PIE-1, five near frontal poses (C27, C05, C29, C09 and C07) and illumination indexed as 08 and 11 were used. The data set was randomly divided into training and test sets; and two samples per person was used for training. We extracted 40 Gabor features. Table II shows the detailed face recognition accuracies. The results clearly demonstrate that MPCA+Improved MDA is superior to all other algorithms. As we knew, this database is really hard for algorithms and most of them had a problem with that. As we can see, our algorithms perform a really good job here and have the most accuracy and also work faster than the others, especially the Improved algorithm, MPCA+Improved MDA, because it eliminates the extra time that we should spend to find the best dimension.

TABLE 2: Recognition Accuracy (%) Comparison of Eigenface, Fisherface, MDA, MPCA+MDA, of MPCA+ Improved MDA with tensors of different orders on PIE-1 Database

Algorithms	**Accuracy**
Eigenface (Grey)	57.2
Eigenface (Gabor)	70.5
Fisherface (Grey)	67.9
Fisherface (Gabor)	76
MDA/2-1 (Grey)	72.9
MDA/2-2 (Grey)	80.4
MDA/3-3 (Gabor)	83.6
MPCA+MDA	87.2
MPCA + Improved MDA	87.5

Another sub-database PIE-2 consists of the same five poses as in PIE-1, but the illumination indexed as 10 and 13 were also used. Therefore, the PIE-2 database is more difficult for classification. We conducted three sets of experiments on this sub-database. As we can see in Table 3, in all the three experiments, MPCA + Improved MDA performs the best and the eigenface has the worst performance. Especially in the cases with a small number of training samples. Also for gaining that performance from our algorgorithm we don't have to spend much time that we use for MPCA+MDA and because of that privilege, our algorithm became a great algorithm to choose.

TABLE 3: Recognition Accuracy (%) Comparison of MPCA+ Improved MDA, MPCA+MDA, Eigenface, Fisherface, MDA/2-1 and MDA/2-2 on the PIE-2 Database

Algorithms	**Test-Train**		
	4-6	**3-7**	***4-6***
Eigenface	39.3	Eigenface	39.3
Fisherface	79.9	Fisherface	79.9
MDA/2-1	74.1	MDA/2-1	74.1
MDA/2-2	81.9	MDA/2-2	81.9
MPCA + MDA	84.1	MPCA + MDA	84.1
MPCA + Improved MDA	84.5	MPCA + Improved MDA	84.5

5. CONCLUSION

In this paper, we improve the performance of MPCA + MDA algorithm by optimizing the subspaces dimension and full projection. Full projection is utilized for initialization the changed SMT and the changed SMT is used to find the optimal subspaces dimension. After that, MDA has been applied for supervised dimensionality reduction. Compared with traditional algorithms, such as PCA and LDA, our proposed algorithm effectively avoids the curse of dimensionality dilemma and overcome the small sample size problem and the advantage of this work is finding the subspaces dimension Because in MDA algorithm the number of possible subspace dimensions for tensor objects is extremely high, comprehensive testing for determination of parameters is not feasible so with this work we save that amount of time. We are eager to apply this algorithm for video-based (fourth order tensor) face recognition and we want to explore this work in our future researches.

REFERENCES

[1]. M. Vasilescu and D. Terzopoulos, “Multilinear subspace analysis for image ensembles,” in Proc. Computer Vision and Pattern Recognition, Madison, WI, Jun. 2003, vol. 2, pp. 93–99.

[2]. G. Shakhnarovich and B. Moghaddam, “Face recognition in subspaces,” in Handbook of Face Recognition, S. Z. Li and A. K. Jain, Eds. New York: Springer-Verlag, 2004, pp. 141–168.

[3]. K. Fukunaga, Statistical Pattern Recognition. New York: Academic, 1990.

[4]. P. Belhumeur, J. Hespanha, and D. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection,” IEEE Trans. Pattern Anal. Mach., vol. 19, no. 7, pp. 711–720, Jul. 1997.

[5]. J. Yang, D. Zhang, A. Frangi, and J. Yang, “Two-dimensional PCA: A new approach to appearance-based face representation and recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 1, pp. 131–137, Jan. 2004.

[6]. H. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, “MPCA: Multilinear Principal Component Analysis of Tensor Objects ,” IEEE Trans. Neural Networks, no. 1, vol. 19, pp. 18-39 Aug. 2008.

[7]. L. D. Lathauwer, B. D. Moor, and J. Vandewalle, “A multilinear singular value decomposition,” SIAM J. Matrix Anal. Appl., vol. 21, no. 4, pp. 1253–1278, 2000.

[8]. S. Yan, D. Xu, Q. Yang, L. Zhang, X. Tang, and H.-J. Zhang, “Multilinear discriminant analysis for face recognition,” IEEE Trans. Image Process., vol. 16, no. 1, pp. 212–220, Jan. 2007.

[9]. A. Shams baboli and G. Rezai-rad, “MPCA+MDA: A Novel approach for Face Recognition Based on Tensor,” presented at the 5th Int. Symposium on Telecommunication, Tehan, Iran, 2010.

[10]. The Olivetti & Oracle Research Laboratory Face Database of Faces, 2002.

[11]. T. Sim, S. Baker, and M. Bsat. “The CMU Pose, Illumination, and Expression (PIE) Database”, Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition, May, 2002.

Tracking Chessboard Corners Using Projective Transformation for Augmented Reality

S. Malek s_malek@cdta.dz
Advanced Technologies Development
Center (CDTA)
Algiers, 16005. Algeria

Zenati-Henda nzenati@cdta.dz
Advanced Technologies Development
Center (CDTA)
Algiers, 16005. Algeria

M.Belhocine mbelhocine@cdta.dz
Advanced Technologies Development
Center (CDTA)
Algiers, 16005. Algeria

S.Benbelkacem sbenbelkacem@cdta.dz
Advanced Technologies Development
Center (CDTA)
Algiers, 16005. Algeria

Abstract

Augmented reality has been a topic of intense research for several years for many applications. It consists of inserting a virtual object into a real scene. The virtual object must be accurately positioned in a desired place. Some measurements (calibration) are thus required and a set of correspondences between points on the calibration target and the camera images must be found. In this paper, we present a tracking technique based on both detection of Chessboard corners and a least squares method; the objective is to estimate the perspective transformation matrix for the current view of the camera. This technique does not require any information or computation of the camera parameters; it can used in real time without any initialization and the user can change the camera focal without any fear of losing alignment between real and virtual object.

Keywords: Pinhole Model, Least Squares Method, Augmented Reality, Chessboard Corners Detection.

1. INTRODUCTION

The term Augmented Reality (AR) is used to describe systems that blend computer generated virtual objects with real environments. In real-time systems, AR is a result of mixing live video from a camera with computer-generated graphical objects that are recorded in a user's three-dimensional environment [1]. This augmentation may include labels (text), 3D rendered models, or shading and illumination variations.
In order for AR to be effective, the real and computer-generated objects must be accurately positioned relative to each other. This implies that certain measurements or calibrations need to be made initially.

Camera calibration is the first step toward computational computer vision. The process of camera calibration determines the intrinsic parameters (internal characteristics of camera) and/or extrinsic parameters of the camera (camera position related to a fixed 3D frame) from correspondences between points in the real world (3D) and their projection points (2D) in one or more images.

There are different methods used to estimate the parameters of the camera model. They are classified in three groups of techniques:

* *Non linear optimization techniques*: the camera parameters are obtained through iteration with the constraint of minimizing a determined function. These techniques are used in many works [2][3]. There advantage is that almost any model can be calibrated and accuracy usually increases by increasing the number of iterations. However, these techniques require a good initial guess in order to guarantee convergence. Surf

* *Linear techniques which compute the transformation matrix*: due to the slowness and computational burden of the first techniques, closed-form solutions have also been suggested. These techniques use the least squares method to obtain a transformation matrix which relates 3D points with their projections [4][5].

* *Two-steps techniques*: these approaches consider a linear estimation of some parameters while others are iteratively estimated.

To ensure tracking of the virtual object, the camera position must be estimated for each view. Several techniques can be used. They are classified in two groups: vision-based tracking techniques and sensor-based tracking techniques [6].

Most of the available vision-based tracking techniques are divided into two classes: feature-based and model-based. The principle of the feature-based methods is to find a correspondence between 2D image features and their 3D world frame coordinates. The camera pose can then be found from projecting the 3D coordinates of the feature into the observed 2D image coordinates and minimizing the distance to their corresponding 2D features [7]. The features extracted are often used to construct models and use them in model-based methods; edges are often the most used features as they are computationally efficient to find and robust to changes in lighting [6].
The sensor-based techniques are based on measurements provided by different types of sensors: magnetic, acoustic, inertial, GPS ... In augmented reality, they are mainly combined with techniques of the first group (hybrid tracking). These methods have the advantage of being robust to abrupt movements of the camera, but they have the disadvantage of being sensitive to environmental disturbances or a limited range in a small volume.

Recently, there is another classification more used to distinguish between vision-based tracking techniques: fiducial marker tracking and markerless tracking [8]. In the first one, the fiducial marker is surrounded by a black rectangle or circle shape boundary for easy detection. Markerless techniques are the other vision-based tracking techniques which don't use fiducial marker.

The common thing between augmented reality applications is the necessity to calibrate the camera at first; if the user wants to change the camera, the resolution or the focal length (zoom lens), he must then calibrate his camera again before starting working with it.

This paper introduces a technique to model a camera using a robust method to find the correspondences without any need of calibration. We use the least squares method to estimate the Perspective Transformation Matrix. For tracking, we use chessboard corners as features to track; these corners are extracted from each video image and used to estimate the Perspective Transformation Matrix. This system does not need any camera parameters information (intrinsic and extrinsic parameters which are included in the Perspective Transformation Matrix). The only requirement is the ability to track the chessboard corners across video images.

The rest of the paper is organized as follows. Section 2 describes related work. Section 3 discusses and evaluates the proposed approach of tracking using a chessboard pattern. In section 4, we present how virtual objects are inserted into real scenes and we give some results. Finally, in the last section, conclusions and discuss of results are given.

2. RELATED WORKS

There are many works which use fiducial marker for tracking. ARToolKit [9][10] is one of the most popular library for augmented reality, it use a rectangular shape which contains an arbitrary grayscale patterns. When the marker is detected, its pattern is extracted and cross-correlated with all known patterns stocked on its data-bases. the camera pose is estimated by using the projective relation between the vertices of the fiducial marker in the scene and the world. ArtoolKit has the inconvenient to be more slowly when using several pattern and markers. However, later research suggested to apply digital communication coding techniques to improve the system's performance, at the cost of customization: QR-code [11], TRIP [12], Artag [13], Cantag [14].

In markerless tracking, the principle is to match features extracted from scenes taken from different viewpoints. Lowe [15] presents a new method named SIFT (Scale Invariant Feature Transform) for extracting distinctive invariant features from images. These features can be used to perform reliable matching between different viewpoints; they are invariant to image scale and rotation, and present robust matching when changing in illumination or adding of noise. Bay et al. [16] present a novel scale- and rotation-invariant interest point detector and descriptor, named SURF (Speeded Up Robust Features). It approximates and also outperforms previously proposed methods to get more robustness and to be much faster.

Recently, Lee and Höllerer [17] extract distinctive image features of the scene and track them frame-by-frame by computing optical flow. To avoid problem of consuming time of processing and achieve real-time performance, multiple operations are processed in a synchronized multithreaded manner.

3. IMPLEMENTATION

We will present here the camera model and the tracking method used for augmented reality.

3.1 Camera Model

The model is a mathematical formulation which approximates the behavior of any physical device, e.g. a camera. There are several camera models depending on the desired accuracy. The simplest is the pinhole Model.

In an AR system, it is necessary to know the relationship between the 3D object coordinates and the image coordinates. The pinhole model defines the basic projective imaging geometry with which 3D objects are projected onto a 2D image plane.
This is an ideal model commonly used in computer graphics and computer vision to capture the imaging geometry. The transformation that maps the 3D world points into the 2D image coordinates is characterized by the next expression:

$$\begin{pmatrix} su \\ sv \\ s \end{pmatrix} = \begin{pmatrix} \alpha_u & \gamma & u_0 & 1 \\ 0 & \alpha_v & v_0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \qquad (1)$$

Equation (1) can be simplified to:

$$s \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = A.R.t \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \qquad (2)$$

With: $[u,v,1]^T$ represents a 2D point position in Pixel coordinates. $[X,Y,Z]^T$ represents a 3D point position in World coordinates. t describes the translation between the two frames (camera frame and world frame), and R is a 3x3 orthonormal rotation matrix which can be defined by the three Euler angles and A is the intrinsic matrix containing 5 intrinsic parameters.
The product *A.R.t* represents the “Perspective Transformation Matrix” M, with:

$$M = A.R.t = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{pmatrix} \qquad (3)$$

From equations (2) and (3), we get:

$$\begin{cases} u = \dfrac{m_{11}X + m_{12}Y + m_{13}Z + m_{14}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \\ v = \dfrac{m_{21}X + m_{22}Y + m_{23}Z + m_{24}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \end{cases} \qquad (4)$$

Each 3D point gives two equations. Six points are then sufficient to estimate the 12 coefficients of the matrix M. But it is possible to use more than 6 points to get better precision. To calculate the different parameters, the constraint $m_{34} = 1$ is used.
To solve the system (4), we first transform it into a linear system as described by (5):

$$\begin{cases} u_1 = m_{11}X_1 + m_{12}Y_1 + m_{13}Z_1 + m_{14} - m_{31}X_1.u_1 - m_{32}Y_1.u_1 - m_{33}Z_1.u_1 \\ v_1 = m_{21}X_1 + m_{22}Y_1 + m_{23}Z_1 + m_{24} - m_{31}X_1.v_1 - m_{32}Y_1.v_1 - m_{33}Z_1.v_1 \\ . \\ . \\ . \\ u_N = m_{11}X_N + m_{12}Y_N + m_{13}Z_N + m_{14} - m_{31}X_N.u_N - m_{32}Y_N.u_N - m_{33}Z_N.u_N \\ v_N = m_{21}X_N + m_{22}Y_N + m_{23}Z_N + m_{24} - m_{31}X_N.v_N - m_{32}Y_N.v_N - m_{33}Z_N.v_N \end{cases} \qquad (5)$$

Then, we transform this system in a matrix form:

$$U = P.V_m \text{ (eq 6)}$$

$$\begin{pmatrix} u_1 \\ v_1 \\ \\ . \\ \\ . \\ \\ . \\ \\ u_N \\ v_N \end{pmatrix} = \begin{pmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & X_1 & Y_1 & Z_1 \\ \\ . & . & . & . & . & . & . & . & . & . & . \\ \\ . & . & . & . & . & . & . & . & . & . & . \\ \\ . & . & . & . & . & . & . & . & . & . & . \\ \\ X_6 & Y_6 & Z_6 & 1 & 0 & 0 & 0 & 0 & X_N & Y_N & Z_N \\ 0 & 0 & 0 & 0 & X_N & Y_N & Z_N & 1 & X_N & Y_N & Z_N \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \end{pmatrix} \qquad (6)$$

To find the m_{ij} parameters, the least squares method is used. The following relation is obtained:

$$V_m = (P^T.P)^{-1}.\ P^T.U \qquad (7)$$

Equation (7) represents a system of equations which can be solved by using a numerical technique such as the Gauss-Jacobi technique.

3.2 Checkpoints Auto Detection and Tracking

We use chessboard corners as features to locate the position of virtual objects in the world coordinate; these corners are extracted from each video frame (Webcam) and used to compute the perspective transformation matrix for the current view. The popular open source library for computer vision OpenCV [18] is used. It provides a function for automatically finding grids from chessboard patterns (*cvFindChessboardCorners()* and *cvFindCornerSubPix()*).

To ensure good tracking and detection, some precautions must be taken which are summarized in the following points:

- The function cvFindChessboardCorners() lists the detected corners line-by-line from left to right and bottom to top according to the first detected corner. There thus are four possibilities of classification which produce errors in the localization of the 3D world reference. In order to avoid this, a rectangular chessboard (it contains MxN corners with M≠N) is used instead of a square one and the corner where the square on its top/right direction is white is considered as the first detected corner. The center of the 3D world coordinate is fixed on the first corner of the second line (Figure1).

- The two faces of the 3D object which contain the chessboard are inclined instead of being perpendicular (Figure1); this configuration facilitates the detection of the checkpoints (chessboard corners) by minimizing the shadow effects and also gives the camera more possibilities of moving.

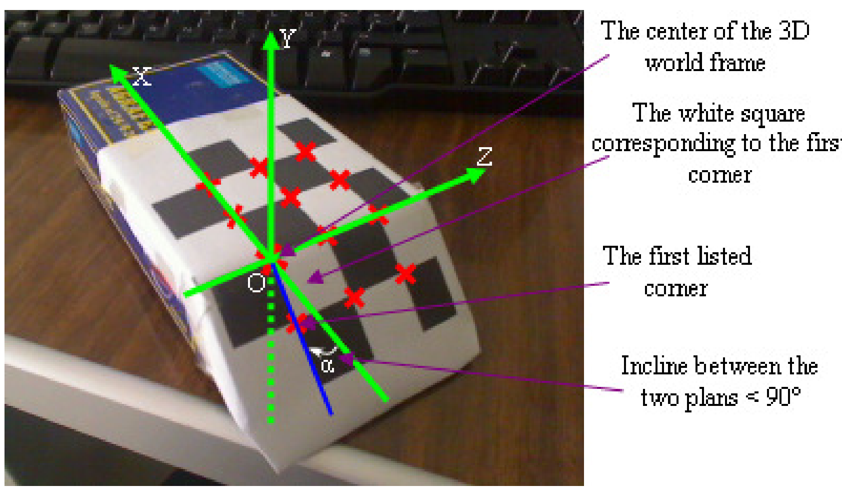


FIGURE1: Example of a 3D object with a 3x4 chessboard

3.3 The Performance Evaluation

To evaluate our method, several measures are performed using three cameras. All tests are performed on the rate frame of 20 Hz with the resolution of 640x480.

The quality of each measurement is estimated by computing the re-projection error, which is the Euclidean distance between the identified corner and the projection of its corresponding 3D coordinates onto the image; about 100 calculations are performed for each case. The obtained results, which represent the average of each case, are represented on the Table1 and compared with results obtained by Fiala and Shu [19].

Our camera model method		Results obtained by Fiala and Shu	
Camera	**Reproj.error (pixel) Std.Dev/Max**	**Camera**	**Reproj.error (pixel) Std.Dev/Max**
Sony VGP-VCC3 0.265	0.14/0.57	**Creative webcam live ultra**	0.16/1.97
Logitech QuickCam Pro 9000	0.15/0.53	**SONY 999 NTSC camera**	0.13/1.25
VRmUsbCam	0.18/0.62	**Logitech QuickCam Pro 4000**	0.29/2.62

TABLE1: Comparison of our camera model method with results obtained by Fiala and Shu [19]

According to Table1, we note that our obtained results are very close from those obtained by Fiala and Shu.

The last important factor of this tracking technique is its robustness against the problem of lighting change; actually corners have the propriety to be detected in different lighting condition. Figure 2 shows an example of corners detection in a very low lighting environment and Figure 3 shows another example in a very high lighting environment.

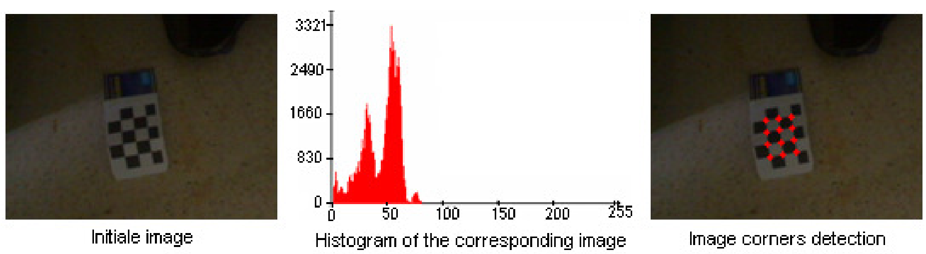


FIGURE 2: corners detection in a very low lighting environment

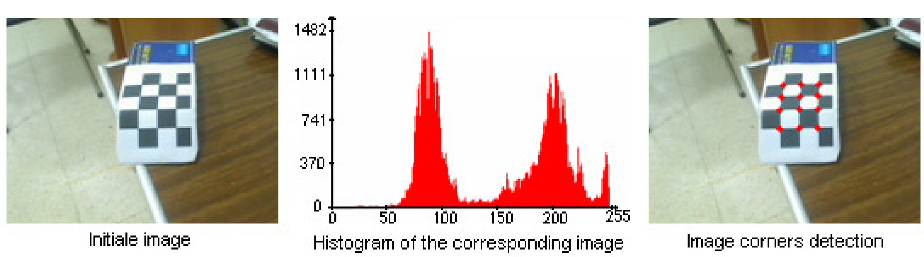


FIGURE 3: corners detection in a very high lighting environment

4. AUGMENTATION TECHNIQUE AND RESULTS

4.1 Principle of Virtual Object Insertion

With the proposal technique, the extrinsic parameters are not required to be calculated. We draw the virtual object directly on the current image using the standard graphic libraries provided by the C++Builder environment.

The principle of insertion is simple; equation (4) is used to compute the projection of any 3D point on the current image. To insert a segment, a straight line is drawn between the projections of its two extremities. For a polygon, it can be inserted using the projection of its apexes. The case of a 3D object is somehow more complicated; this object can be inserted on the current image by drawing its faces one by one. An example of a cube is presented in figure4. In this example, a virtual cube is drawn and positioned as mentioned in figure5. At first, the projections of its apexes (P_i : i = 1..8) are calculated, then its six faces are drawn one by one in a respective order which is described in the next sub-section.

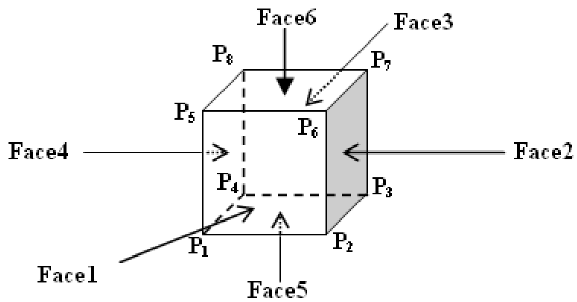


FIGURE4: the virtual object to insert (cube)

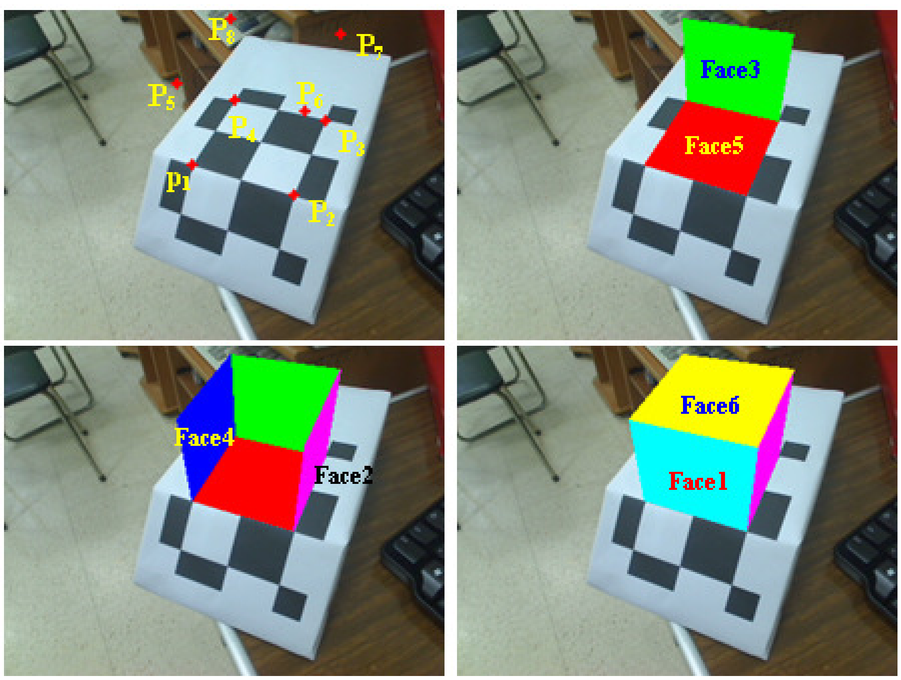


FIGURE5: The technique used to insert a 3D object

4.2 Technique Used to Avoid Occultation Problem

Occultation problem between real and virtual objects are not treated here, but only the occultation between the virtual object components, i.e. the different faces which form the augmented object. The virtual cube shown in figure5 is inserted in a special order (Face5, Face3, Face4, Face2, Face1 and Face6). If this order changes, the inserted object does not appear as a cube. Figure6 depicts a wrong presentation of a virtual object (house) when its faces are inserted in another order; the front faces are occulted by others since they are inserted in advance.

We note that this technique is not necessary when using the wire frame technique for virtual object insertion.

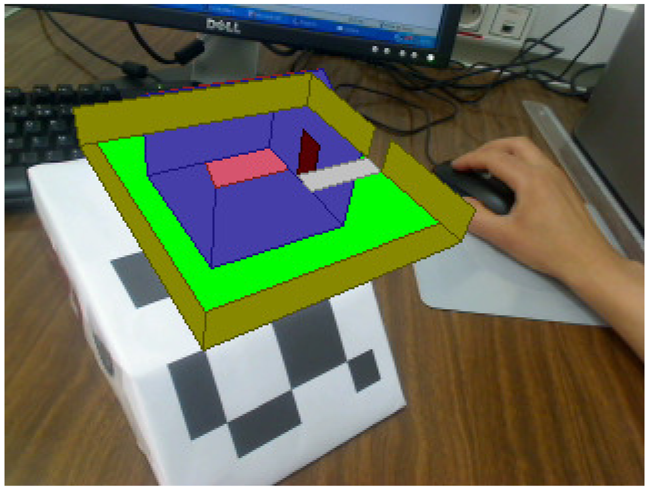


FIGURE6: Example of wrong augmentation

This problem can be avoided by estimating the camera viewpoint. This viewpoint consists of the estimation of the camera orientation according to the direction of the 3D world reference axes. We note that this technique is not required if the virtual object is drawn by the wire frame technique.

The technique is based on the use of a virtual Cube put on the horizontal plan (OXZ) where its vertices are $S_1(0,0,0)$, $S_2(0,0,L)$, $S_3(L,0,L)$, $S_4(L,0,0)$, $S_5(0,L,0)$, $S_6(0,L,L)$, $S_7(L,L,L)$ and $S_8(L,L,0)$. L can take any positive value. Projections of these 3D points on the current image ($s1(u_1,v_1)$, $s_2(u_2,v_2)$, $s_3(u_3,v_3)$, $s_4(u_4,v_4)$, $s_5(u_5,v_5)$, $s_6(u_6,v_6)$, $s_7(u_7,v_7)$, $s_8(u_8,v_8)$) are then calculated using equation 5 (see figure7 and figure8). Once calculated, a set of tests are applied using these projections to estimate the camera viewpoint.

These tests are divided into two groups; in the first group, we use the four vertices of the bottom-square to estimate their positions according to the camera. For the second one, we need to calculate the projection of "s_5" on the lines (s_1s_2) and (s_1s_4) and the projection of "s_7" on the lines (s_2s_3) and (s_3s_4); these projections are respectively $p_{5\_12}(u_{5\_12},v_{5\_12})$, $p_{5\_14}(u_{5\_14},v_{5\_14})$, $p_{7\_32}(u_{7\_32},v_{7\_32})$ and $p_{7\_34}(u_{7\_34},v_{7\_34})$. Then, using these projections we estimate the faces which are opposed to the camera.

In the example shown in figure7, the verified tests are:

- ($u_4 < u_1 < u_2$) and ($v_5 > v_{5\_12}$) and ($v_5 > v_{5\_14}$) which means that the camera is oriented according to $\overrightarrow{OX}$ and $\overrightarrow{OZ}$ axes.

In the example shown in figure8, the verified tests are:

- ($u_1 < u_2 < u_3$) and ($v_5 > v_{5\_12}$) and ($v_7 > v_{7\_32}$) which means that the camera is oriented according to $\overrightarrow{OX}$ and $-\overrightarrow{OZ}$ axes.

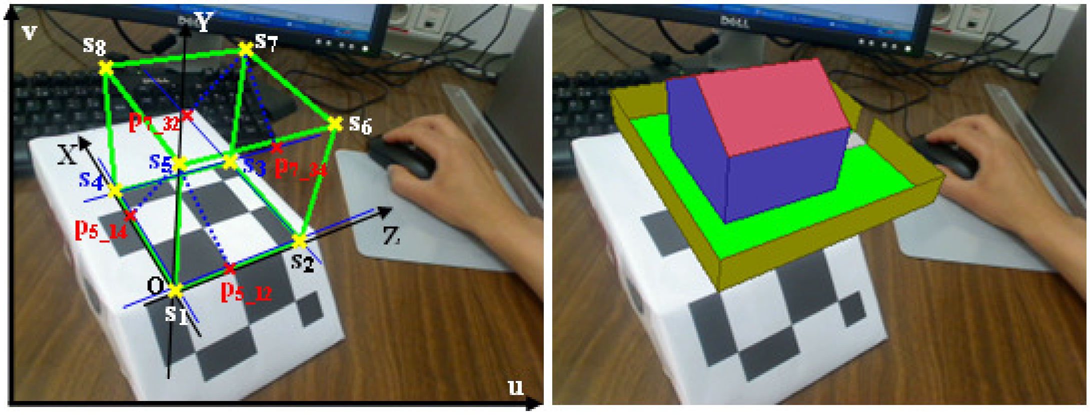


FIGURE7: Example 1 of the camera estimation viewpoint

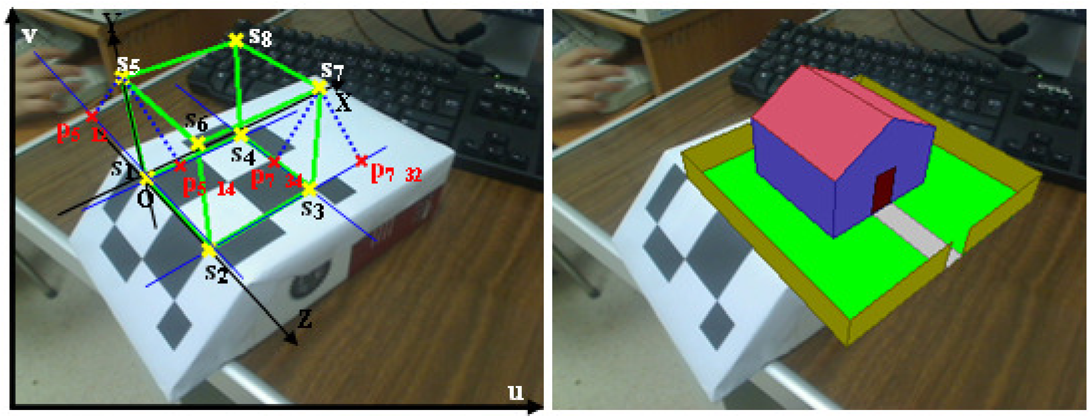


FIGURE8: Example 2 of the camera estimation viewpoint

4.3 Results

We present here two application examples of our approach. The scene captured by the camera (webcam) is augmented by virtual objects in real time. In the first example, the "Logitech QuickCam Pro 9000" webcam is used with the resolution 320x240, in the second one; a "Sony VGP-VCC3 0.265A" webcam is used with a resolution 640x480. Images presented in figure9 and figure10 are selected arbitrarily from the webcams.

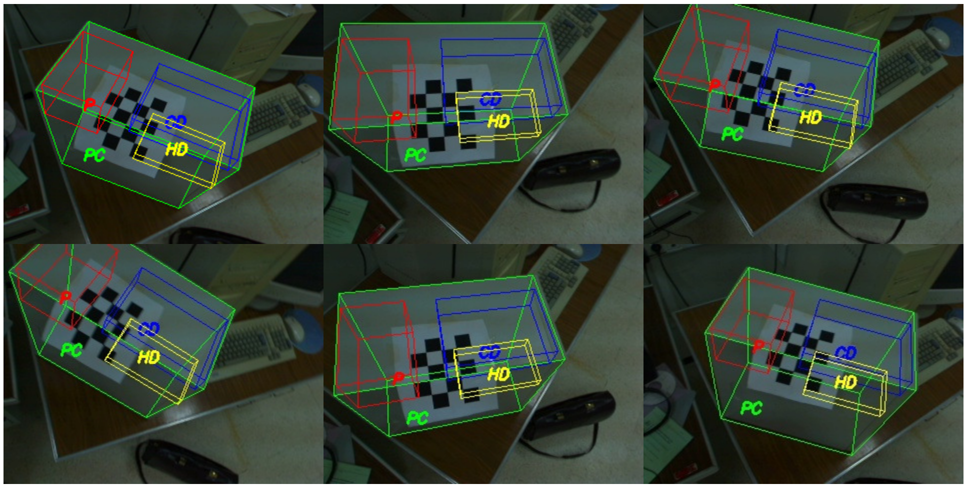


FIGURE9: Example of virtual object insertion into real scene with the wire frame technique

FIGURE10: Example of augmentation with a 3D virtual model.

According to these examples, we note that using 12 fiducial points (corners) is sufficient to model the camera using a pinhole model; the virtual object is positioned in the desired place with a good accuracy and it size depends of the camera depth. We also note that the used technique of tracking (tracking of chessboard corners) gives good and acceptable results and the virtual object is well tracked across camera frame in real time. The last point is the efficacy of the developed technique of virtual object insertion; the virtual object follows the motion of the camera and it is positioned and drawn correctly without any information about intrinsic or extrinsic parameters of the camera.

5. CONCLUSION

This paper introduces a technique of virtual object tracking in real time. This technique is developed as part of an augmented reality system and does not require any information or computation of the camera parameters (intrinsic and extrinsic parameters). It is based on both detection of Chessboard corners and a least squares method to estimate the perspective transformation matrix for the current view of the camera.

The use of this technique is simple and does not require initialization steps or manual intervention, the only requirement is the tracking of the marker (chessboard) across images provided by the camera.

The results show the efficiency of the tracking method based on detection of chessboard corners; the major advantages of tracking corners are their detection robustness at a large range of distances, their reliability under severe orientations, and their tolerance of lighting changes or shadows. The technique used for the insertion of virtual objects gives its proof when the camera parameters are not calculated.

6. REFERENCES

1. R. Azuma, Y. Baillot,R. Behringer, S. Feiner, and S. Julier, "*Recent advances in augmented reality*", IEEE Computer Graphics and Applications, 21(6), pp. 34-47, 2001.

2. R. Tsai, "*An efficient and accurate camera calibration technique for 3D machine vision*", IEEE Proceedings CVPR, pp. 364-374, June 1986.

3. M. Tuceryan, M., Greer, D., Whitaker, R., Breen, D., Crampton, C., Rose, E., and Ahlers, K., "*Calibration Requirements and Procedures for Augmented Reality*", IEEE Trans. On Visualization and Computer Graphics, pp. 255-273, September 1995.

4. O.D. Faugeras and G. Toscani. "*The calibration problem for stereo*", In Proceedings of CVPR, pp. 15-20, 1986.

5. Z. Zhang, "*Camera calibration with one-dimensional objects*", In Proc. 7th European Conference on Computer Vision, (IV), pp. 161–174, 2002.

6. Zhou, F., Duh, H.B.L., Billinghurst, M. " *Trends in Augmented Reality Tracking, Interaction and Display: A Review of Ten Years of ISMAR*". Cambridge, UK: 7th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2008), pp. 15-18, Sep 2008.

7. H. Wuest, F. Vial and D. stricker. "*Adaptative line tracking with multiple hypotheses for augmented reality*". In ISMAR '05, pp. 62-69, 2005.

8. Hyun Seung Yang, Kyusung Cho, Jaemin Soh, Jinki Jung, Junseok Lee: "*Hybrid Visual Tracking for Augmented Books"*. ICEC 2008: pp. 161-166, 2008.

9. H. Kato and M. Billinghurst. "*Marker tracking and HMD calibration for a video-based augmented reality conferencing system"*. IWAR '99 : Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality, Washington, DC, USA. IEEE Computer Society, pp. 85–92, 1999.

10. ARToolKit site: http://www.hitl.washington.edu/artoolkit.

11. ISO. International standard, iso/iec18004. ISO International Standard, Jun 2000.

12. Diego Lopez de Ipina, Paulo R. S. Mendonca, and Andy Hopper. "*Trip: A low-cost vision-based location system for ubiquitous computing"*. Personal Ubiquitous Comput., 6(3), pp. 206–219, , 2005.

13. Fiala, M.: ARTag: *"a fiducial marker system using digital techniques*". In: CVPR'05, vol. 1, pp. 590 – 596, 2005.

14. Rice, A., Beresford, A., Harle, R.: "*Cantag: an open source software toolkit for designing and deploying marker-based vision systems"*. In: 4th IEEE International Conference on Pervasive Computing and Communications, 2006.

15. Lowe, D.G.: "*Distinctive Image Features from Scale-Invariant Keypoints*". International Journal of Computer Vision 20(2), pp. 91–110, 2004.

16. Bay, H., Tuytelaars, T., VanGool, L.: "*SURF: Speeded Up Robust Features*". In: 9th European Conference on Computer Vision, pp. 404–417, 2006.

17. T. Lee and T. Höllerer. 2009. "*Multithreaded Hybrid Feature Tracking for Markerless Augmented Reality*". IEEE Transactions on Visualization and Computer Graphics 15, 3, pp. 355-368, May. 2009.

18. Open source computer vision library. In: http://www.intel.com/research/mrl/research/opencv.

19. Mark Fiala, Chang Shu: "*Self-identifying patterns for plane-based camera calibration*". Mach. Vis. Appl. 19(4), pp. 209-216, 2008.

Histogram Gabor Phase Pattern and Adaptive Binning Technique in Feature Selection for Face Verification

Srinivasan A asrini30@gmail.com
Professor
Department of Computer Science and Engineering
Misrimal Navajee Munoth Jain Engineering College
Chennai, Tamilnadu, India.

Abstract

The aim of this paper is to develop a robust system for face recognition by using Histogram Gabor Phase Pattern (HGPP) and adaptive binning technique. Gabor wavelet function is used for representing the features of the image both in frequency and orientation level. The huge feature space created by Gabor wavelet is classified by using adaptive binning technique. The unused bin spaces are used. As a result of which, the size of the space is drastically reduced and high quality HGPP created. It is due to this approach, the computation complexity and the time taken for the process is reduced and the recognition rate of the face improved. The significance of this system is its compatibility in yielding best results in the face recognition with major factors of a face image. The system is verified with FERET database and the results are compared with those of the existing methods.

Keywords: Face Recognition, Gabor Wavelets, Local Gabor Phase pattern, Global Gabor Phase Pattern, Adaptive Binning, and Spatial Histograms.

1. INTRODUCTION

Face recognition, the most coveted field in Image Processing, is still in its initial stages. It is due to its scientific challenges and potential applications, Face recognition has been an active research topic over the past few years. The Gabor wavelets approach appears to be quite perspective and it has several advantages such as invariance to some degree with respect to homogenous illumination changes, small changes in head poise, robustness against facial hair, and image noise [1,2]. Experimental results show that the proposed method performs better than traditional approaches in terms of both efficiency and accuracy.

To eliminate extrinsic factors, various feature extraction and selection methods are used. One such a method is HGPP. In this method, the quadrant bit codes are first extracted from face based on the Gabor transformation and Histogram techniques. The features of HGPP lie in two aspects. They are: i) HGPP can describe the general face images robustly without training procedure, ii) Encodes the Gabor phase Information, instead of Gabor magnitude information. This method uses two Gabor Phase Patterns (GPP's) to encode the phase variations which use high dimensional histogram features resulting in performance decrease and computational complexity. In this proposed system, the above stated problems are rectified by using adaptive binning method. As a result, the overall efficiency of the face recognition system is increased.

1.1. Challenges Associated With Face Verification

Face verification and recognition is a challenging problem due to variations in pose, illumination, and expression. So, Techniques that can provide effective feature representation with enhanced discriminability are crucial [3]. Face recognition has become one of the most active research areas and it plays an important role in many applications such as human machine interaction, authentication, and surveillance. However, the wide range variations of human face due to pose, illumination, and expression not only result in a highly complex distribution but also deteriorate the verification rate. It seems impractical to collect sufficient prototype images covering all the possible variations. Therefore, it is highly imperative in research to construct a small size training face verifier which is robust to environmental variations.

2. EXPERIMENTS

2.1. Block Diagram

The Figure 1 shows the entire block design of the proposed system with new methodology.

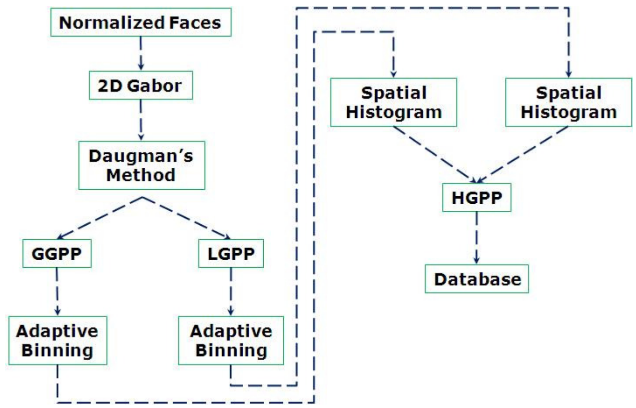


FIGURE 1: Block design for the proposed system

The Normalized face is given as input to the four different processes i) gabor filters ii) daugman's method, iii) Global Gabor Phase Pattern (GGPP) and iv) Local Gabor Phase Pattern (LGPP). For the Adaptive binning, the result obtained from GGPP and LGPP is given as input. This method works by creating a bin of size 3x3. And the resultant value obtained from adaptive binning is given to Spatial Histogram and the outputs of spatial histogram are used to create HGPP. Using HGPP value for the test image and trained dataset the relevant matched images are obtained.

Adaptive binning is a method for binning images according to the local count rate. It attempts to adaptively bin a single image, based on the number of pixels in each region. The basic method is to bin pixels in two dimensions by a factor of two, until the fractional Poisson error of the count in each bin becomes less than or equal to a threshold value. When the error is below this value, those pixels are not binned any further. The algorithm starts with the smallest possible bin size of 1×1 pixel and then calculates the average mean count for each bin. Each bin with an average mean count higher than the threshold value is marked as binned and its pixel members are removed from the pixel list, and ignored during the rest of the binning process. In the next iteration, the unbinned pixels are rebinned with square bins of double the side length. This process is repeated until all pixels are binned or the bin size exceeds the image size. Adaptive Binning of image feature is done, since the number of features obtained after the GGP operations is large. In order to reduce the number of features and at the same time to retain meaningful information adaptive binning is performed.

2.2. Gabor Wavelets

Gabor feature has been recognized as one of the best representations for face recognition. Traditionally, only the magnitudes of the Gabor coefficients are thought to be valuable for face recognition, and the phases of Gabor features are deemed to be useless and always discarded directly by almost all researchers in face recognition community [4, 5]. When The spatial histograms generated by encoding Gabor phases through Local Binary Pattern (LBP) they yield better recognition rate comparable with that of Gabor magnitude based methods. And it is also shown that the Gabor phases are quite compensatory to the magnitude information, since higher classification accuracy is achieved by combining Gabor phases and magnitudes. All these observations suggest that more attention should be paid to Gabor phases for face recognition. Among various wavelet bases, Gabor functions provide the optimized resolution both in the spatial and frequency domains.

2.3. Gabor Wavelets Functions

Gabor wavelet are obtained by using eq-1

$$\Psi_{u,v}(z) = \frac{(\|K_{u,v}\|)^2}{\sigma^2} e^{\frac{(-\|K_{u,v}\|z\|)^2}{2\sigma^2}} \left[e^{iK_{u,v}z} - e^{-\frac{\sigma^2}{2}} \right] \qquad \text{.............1}$$

where $K_{u,v} = \begin{pmatrix} K_{ju} \\ K_{jv} \end{pmatrix} = \begin{pmatrix} K_v \cos\theta_u \\ K_v \sin\theta_u \end{pmatrix}$

where $K_v = \frac{f_{max}}{2^{v/2}}, \sigma = 2\pi,\ \theta_u = u\left(\frac{\pi}{8}\right)$ $and\ Z = image\ position$
The equation (1) is again simplified into

$$f_{u,v}(z) = \frac{D}{\sigma^2} e^{\frac{-D.E}{2\sigma^2}} \left[e^F \quad e^{-\frac{\sigma^2}{2}} \right]$$

where $D = \left(\|K_{u,v}\|\right)^2$, $E = (\|Z\|)^2$ and $F = iK_{u,v}Z$. The parameters D & E are calculated using l^2 norms. l^2-norm is defined as the square root of sum of squares of individual components. The l^2-norm is a vector norm defined for a complex vector. The l^2-norm is the vector norm that is commonly used in vector algebra and vector operations (such as the dot product), where it is commonly denoted by$|x|$. However, if desired, a more explicit (but more cumbersome) notation $|x|$ can be used to emphasize the distinction between the vector norm $|x|$ and complex modulus$|z|$. But the fact is that the l^2-norm is just one of the possible type of norms. The l^2 -norm of the vector $x = (x_1, x_2, x_3)$is given by $|X| = \sqrt{x_1^2 + x_2^2 + x_3^2}$ and l^2-norm is represented as$\sqrt{\sum[Abs(A^2)]}$, where A=$max_{i \le j \le m} \sum_{i=1}^{m} |Z_{ij}|$. In the equation (1), '*u*' refers to the orientation and '*v*' refers to frequency, f_{max} is a constant. The equation (1) can be simplified further and one can get real and imaginary parts as shown in equation (2) and (3).

$$R_e(f) = -\frac{A}{\sigma^2} e^{f - \sigma^2/2} \qquad \text{.............2}$$
$$I_{mag}(f) = \frac{A}{\sigma^2} e^{f + icd} \qquad \text{.............3}$$

2.4. Daugman's Method

The real and imaginary parts of the Gabor wavelets are applied to the Daugman's Method proposed by Daugman's for demodulation. When the output of Gabor Wavelets is demodulated, each pixel in the resultant image is encoded to two bits [6]. This method is essential to split the Gabor Wavelets Pattern to GGPP and LGPP [6] and it is done by using the equation (4) and (5).

$$P_{u,v}^{Re}(z) = \begin{Bmatrix} 0\ if\ Re(G_{u,v}(z)) > 0 \\ 1\ \ if\ Re\left(G_{u,v}(z)\right) \le 0 \end{Bmatrix} \qquad \text{.............4}$$
$$P_{u,v}^{Im}(z) = \begin{Bmatrix} 0\ if\ Im(G_{u,v}(z)) > 0 \\ 1\ \ if\ Im\left(G_{u,v}(z)\right) \le 0 \end{Bmatrix} \qquad \text{.............5}$$

where *Re($G_{u,v}$(z))* and *Im($G_{u,v}$(z))* are real and imaginary parts of Gabor coefficient Daugman's encoding method can be reformulated as equations 6 & 7.

$$P_{u,v}^{Re}(z) = \begin{cases} 0\ if\ \theta_{u,v}\ (z) \epsilon\ I, IV \\ 1\ \ if\ \theta_{u,v}(z)\ \epsilon\ II, III \end{cases} \qquad \text{.............6}$$
$$P_{u,v}^{Im}(z) = \begin{cases} 0\ if\ \theta_{u,v}\ (z) \epsilon\ I, IV \\ 1\ \ if\ \theta_{u,v}\ (z) \epsilon\ II, III \end{cases} \qquad \text{.............7}$$

where $\theta_{u,v}(z)$ is the Gabor phase angle for the pixel at the position.

2.5. Formation of GGPP

GGPP scheme computes one binary string for each pixel by concatenating the real and imaginary bit codes of different orientations for a given frequency. Formally, the GGPP value, $GGPP_v(Z_o)$, for the frequency '*v*' at the position Z_o in a given image is formulated as the combination of Daugman's Values. By using this encoding method, a decimal numbers for each pixel corresponding to the real and imaginary GGPPs is obtained. GGPP scheme computes one binary string for each pixel by concatenating the real and imaginary bit codes of different orientations for a given frequency using equation (8) and (9).

$$GGPP_v^{Re}(z_0) = \left[P_{0,v}^{Re}(z_0), P_{1,v}^{Re}(z_0) \dots\dots\dots P_{k,v}^{Re}(z_0)\right] \qquad \text{.............8}$$
$$GGPP_v^{Im}(z_0) = \left[P_{0,v}^{Im}(z_0), P_{1,v}^{Im}(z_0) \dots\dots\dots P_{k,v}^{Im}(z_0)\right] \qquad \text{.............9}$$

The above equations give both real and imaginary GGPP. In this approach there are eight orientations representing 0-255 different orientation modes.

2.6. Formation of LGPP

LGPP is yet another encoding of local variations for each pixel. LGPP actually encodes the sign difference of the central pixel from its neighbors. LGPP reveals the spots and flat area in the given images. Formally, for each orientation '*u*' and frequency '*v*', the real LGPP value at the position Z_o is computed using local XOR pattern (LXP) operator [6, 7]. The local variation for each pixel obtained is LGPP. LGPP actually encodes the sign difference of the central pixel from its neighbors. LGPP reveals the spots and flat area in the given images using equation (10).

$$LGPP_{u,v}(Z_0) = P_{u,v}(Z_0) XOR P_{u,v}(Z_1), P_{u,v}(Z_0) XOR P_{u,v}(Z_2), P_{u,v}(Z_0) XOR P_{u,v}(Z_3), P_{u,v}(Z_0) XOR P_{u,v}(Z_4), P_{u,v}(Z_0) XOR P_{u,v}(Z_5), P_{u,v}(Z_0) XOR P_{u,v}(Z_6), P_{u,v}(Z_0) XOR P_{u,v}(Z_7) \quad \ldots\ldots\ldots\ldots..10$$

where Z_i, i = 1,2,...8 are the eight neighbors around Z_0, and *XOR* denotes the bit exclusive or operator.

3. ADAPTIVE BINNING TECHNIQUES

Histograms are used in image retrieval systems to represent the distributions of colors in images. The histograms adapted to images represent their color distributions more efficiently than histograms with fixed binnings. Adaptive histograms produce good performance, in terms of accuracy, less number of bins and efficient computation when compared to that of the existing methods for retrieval, classification, and clustering tasks. There are two general methods of generating histograms: i) fixed binning and ii) adaptive binning. Adaptive binning is similar to color space clustering in that k-means clustering. In other words, its variant is used to induce the bins. However, the clustering algorithm is applied to the colors in an image instead of the colors in an entire color space. Therefore, adaptive binning produces different bins for different images. The adaptive binning algorithm is applicable to a wide range of data and it is not limited to two dimensional data [8]. Adaptive binning is the simplest case of the algorithm. It attempts to adaptively bin a single image based on the number of pixels in each region. The basic method is to bin pixels in two dimensions by a factor of two, until the fractional Poisson error of the count in each bin becomes less than or equal to a threshold value. When the error is below this value, those pixels are not binned any further [9, 10].

3.1. Adaptive Binning Algorithm

Step 1: Put each pixel in a 'bin', which is a collection of pixels.
Step 2: The net count in the bin is defined by $s_i = c_i - n_i b$
Step 3: Fractional error in the bin is calculated as $\frac{\sigma(s_i)}{s_i} = \frac{\sqrt{c_i + n_i b}}{c_i - n_i b}$
Step 4: Find aveage mean count s_i/n
Fractional error <= threshold value in processed, find average mean count
else bin not yet processed
Step 5: Set Identification number for each processed bin
Step 6: Merge the neighboring bins.
Step 7: Repeat from Step 2 until a single bin is got

4. SPATIAL HISTOGRAMS

Object representation and feature extraction are essential to object detection. Specially, objects are modeled by their spatial histograms over local patches and class specific features are extracted. Spatial histograms consist of marginal distributions of an image over local patches and they can preserve texture and shape information about an object simultaneously [11]. GGPP & LGPP are relatively new and simple texture models proved to be a very powerful feature in classification of images [12]. GGPP & LGPP are invariant against any monotonic transformation of the gray scale and the basic GGPP & LGPP operator uses neighbourhood intensities to calculate the region central pixel value [1]. The 3 x 3 neighbourhood pixels are signed by the value of center pixel using the eq (11)

$$S_{(g_0,g_i)} = \begin{cases} 1, if\ g_i \geq g_0 \\ 0, if\ g_i \geq g_0 \end{cases} \ where\ i \leq i \leq 8 \quad \ldots\ldots\ldots\ldots..11$$

The signs of the eight differences are encoded into an 8-bit number to obtain LGPP value from the center pixel and calculated using eq (12)

$$\sum_{i=1}^{8} S(g_0, g_i)\ 2^{i-1} \quad \ldots\ldots\ldots\ldots..12$$

For any sample image, histogram-based pattern representation is computed as follows, first, variance normalization on the gray image to compensate the effect of different lighting conditions. And then basic global or local binary pattern operator is used to transform the image into an GGPP or LGPP image. Finally, histogram of an image is computed as its representation. It is easy to prove that histogram, a global representation of image pattern, is invariant to translation and rotation. However, histogram technique is not adequate, since it does not encode spatial distribution of objects. For irrelevant and relevant images, their histograms can be very similar or even identical, making histogram insufficient.

After using the spatial histograms and adaptive binning, a new pattern called HGPP is developed. It not only reduces data size but also involves less complexity. As a result, most of the unwanted data are removed and the performance increased. Using this HGPP Patterns, the test images for the specific database is checked and verified.

5. RESULTS AND ANALYSIS

This system uses a normalized image as input. Gabor wavelets, which are directly related to Gabor filter is a linear filter used for edge detection. A set of Gabor filers with different frequencies and orientations are helpful for extracting useful features from an image. In this system, frequency value is set to five and orientation to eight. Gabor filters has a real and an imaginary component representing orthogonal directions. So, in total, 80 (40 real and 40 imaginary) different sets of Gabor filters are obtained from a single image. These filters are further processed to demodulate the image by using Daugman's method. In this process, all the 80 images are demodulated to obtain quantified Gabor feature. After quantifying the Gabor features using Daugman's method, global Gabor phase patterns are generated to form a byte to represent 256 different orientation modes and a total of 10 (5 real and 5 imaginary) images are obtained in GGPP. To encode the local variations in a pixel, local Gabor phase pattern is applied to all 80 Gabor features using local XOR pattern. For five frequency and eight orientations, the phase patterns obtained will be 90 "images" (five real GGPP, five imaginary GGPP, 40 real LGPPs and 40 imaginary LGPPs), with the same size as the original image. To reduce the size of phase patterns, binning is done by using adaptive binning technique. Each phase pattern is taken and binned in such a manner that the count in each bin becomes less than or equal to the threshold value. To reserve the spatial information in the phase patterns, the GPP and LGP images are spatially divided into the non over-lapping rectangular regions, from which the spatial histograms are extracted. Then, all of these histograms are concatenated into a single extended histogram feature called HGPP. Formally, the HGPP feature is formulated as shown in eq (13)

$$HGPP = (H^{Re}_{GGPP}, H^{Im}_{GGPP}, H^{Re}_{LGPP}, H^{Im}_{LGPP}) \quad \ldots\ldots\ldots\ldots.13$$

where $H^{Re}_{GGPP}, H^{Im}_{GGPP}$ are the sub regions of real and imaginary part of GGPP and $H^{Re}_{LGPP}, H^{Im}_{LGPP}$ are the sub regions of real and imaginary part of LGPP. Final HGPP representation is a local model robust to local distortions caused by different imaging factors such as accessory, expression variations.

5.1. Input Image Database

The database is developed into 3 different sizes, such as 64x64, 88x88 and 128x128. Based on the face image factors, images are categorized into 10 parts as shown in Table 1. In each size of image, the images are categorized into 10 parts as shown in Table 1. The images resized in three different sizes are viz. 64x64, 88x88 and 128x128.

Image Factor Type	Description
Aging	Aging subjects
Dup I	Subsets of Aging
Dup II	
Fa	Frontal
Fb	Expression
Fc	Illumination

TABLE 1: Classification of Image Database

5.2. Output - Gabor Wavelets

The images to be trained are given as input to the Gabor functions, where the Gabor wavelets are obtained. The sample Gabor wavelets obtained for images are shown in the figure 2 & 3. In the proposed system, the frequency and orientation of an image are increased by keeping the frequency (v) to five and orientation (u) to eight. For real and imaginary part of a single trained image, the database will have eighty images.

FIGURE 2: Images after applying Gabor Function (Frequency)

FIGURE 3: Images after applying Gabor Function (Orientation)

5.3. HGPP Patterns

Figure 5 and 6 shows the patterns obtained after applying the HGPP and Adaptive binning respectively for the images shown in Figure 4. From the figure, reduction in the size of the data for an image can be observed.

FIGURE 4: Image used

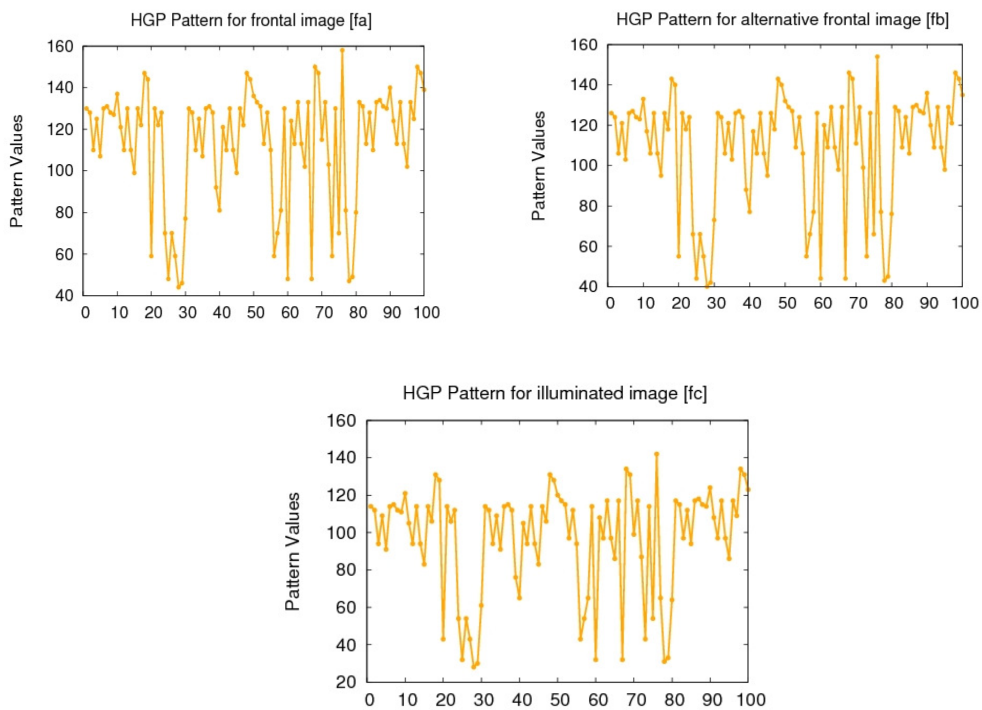


FIGURE 5: HGP Patterns for three image types

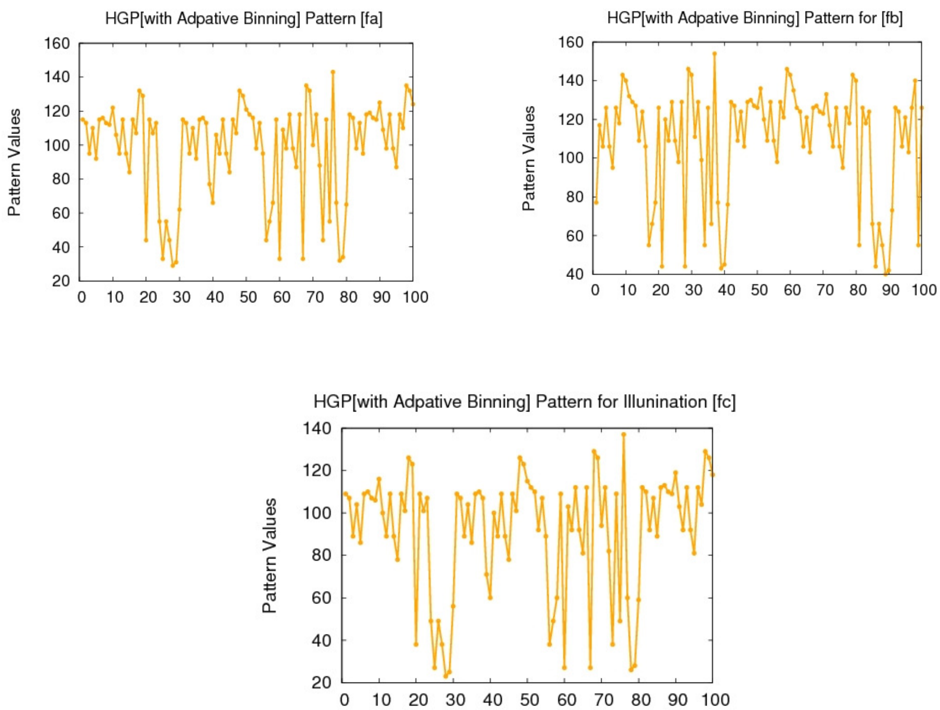


FIGURE 6: HGP Patterns using Adaptive binning for three image types

The recognition rates for different sizes of images are tabulated in the Table 2 and the results plotted are shown as Figure 7. To further validate the effectiveness of HGPP, these results are compared with those available in methods such as Feature Extraction, Eigen Face and HGPP (i.e. without using Adaptive Binning Method). The results clearly indicate that the proposed HGPP method outperforms all the other methods, especially on the Dup I, and Dup II probe set. Experimental results of this comparison evidently prove that the proposed HGPP method achieves best results on our database. Since the face images in our database probe sets contain several source of variations such as expression, lighting, and aging, these comparisons indicate that HGPP is impressively robust to these extrinsic imaging conditions. Gabor features can exhibit the spatial frequency (scale), spatial locality, and orientation selectivity properties corresponding to Gabor wavelets. Adaptive Binning is a kind of quantification of Gabor feature contributing to the robustness of HGPP.

Probe Size	Fa	Fb	Fc	Aging	Dup I	Dup II
64x64, $f_{max} = 3.14$	99.12	98.32	97.48	92.00	89.1	84.11
88x88, $f_{max}=1.11$	99.78	98.74	97.89	93.00	88.66	83.62
128x128, $f_{max}=1.57$	98.34	99.58	99.16	91.00	89.51	85.71

TABLE 2: Recognition rates in percentage for different sizes of image data base.

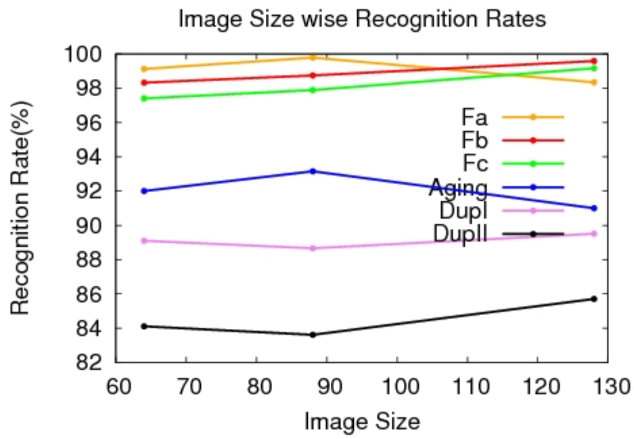


FIGURE 7: Recognition rate for different sizes of images and for probe sets (HGPP)

It can be seen From the Table 2 that when the image is 64x64 better recognition rates are obtained for frontal images [Fa] then for alternative frontal images [Fb]. For illuminated images [Fc] the recognition rates decrease to some extent when compared with the Fa and Fb. For aging [Dup1], one gets less recognition rate value because of the resolution change in the image. f_{max} plays a vital role for getting better recognition rate, when size of image increases and f_{max} value decreases, recognition rate is improved. But in case of 128x128 sized images, f_{max} value has to be increased to get the better efficiency. Hence, fixing the appropriate value of f_{max} plays an important role to get improved performance rate.

Figure 7 shows the graph version of Table 2. The graph is drawn with image size as x-axis and recognition rate as y-axis. It can be seen from the graph that when the image size increases, there is a linear increase in recognition rate, when the image size increases, the clarity of image increases. As a result, the efficiency increases and the time taken to process the image also increases.

Methods	Image Database Probe Set					
	Fa	Fb	Fc	Aging	Dup I	Dup II
Feature Extraction [FE] Method	93.15	93.25	94.56	83.75	85.23	81.25
Eigenfaces [EF] Method	95.78	95.89	95.12	85.12	87.21	82.56
HGPP [HGP] Method	98.00	96.13	96.89	91.75	88.85	83.65
HGPP(with Adaptive Binning) [ADP] Method	99.00	98.32	97.48	92.15	89.10	84.11

TABLE 3: Recognition Rate Comparison for various methods

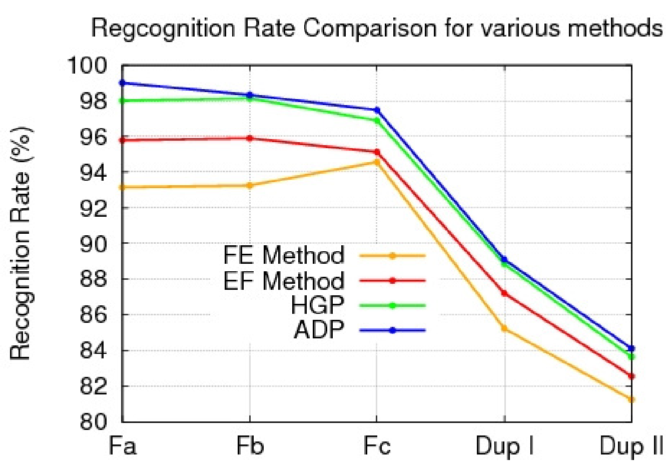


FIGURE 8: Recognition Rate Comparison for various methods

Table 3 presents the efficiency rate comparison for various methods such as Feature Extraction, Eigen Face, HGPP, and HGPP with Adaptive Binning methods. The comparison is done with face image factors such as frontal image [Fa], alternative frontal image [Fb], illuminated image [Fc], and aging images [Dup I and Dup II]. Transforming the input data into set of features is called *feature extraction*. If the features extracted are carefully chosen it is expected that the feature set will extract the relevant information from the input data. Using this reduced representation desired task is performed instead of the full size input. Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately because a number of variables are involved in simplifying the data. Analysis with a large number of variables generally requires a large amount of memory and computation power. It is because of more memory and computation power, the recognition rate efficiency goes down.

In Eigenfaces, a set of eigenfaces are generated by performing a mathematical process called principal component analysis on a large set of images depicting different human faces. Informally, eigenfaces can be considered a set of "standardized face ingredients", derived from statistical analysis of many pictures of faces. And also, it does not take many eigenfaces combined together to achieve a fair approximation of most faces. Each eigenface represents only some features of the face which may or may not be present in the original image. If the feature is present in the original image, then the contribution of that eigenface in the sum of eigenfaces will be greater. Otherwise it achieves a very low approximation of faces. It is because of data loss, the recognition rate efficiency goes down and because of less computation cost and memory, the rate is more than that of the Feature Extraction method.

As regards HGPP, the normalized image is given as input to the Gabor wavelets, from where a lot of processed images are obtained, and these images are given as input to Daugman's and LGP, and GGP patterns are generated and processed with spatial histogram. In this process, HGP patterns are obtained involving a huge amount of data. It is because of this, the time taken for processing and the computational cost are increased. Whereas in Adaptive Binning concept, the pattern obtained after HGPP are binned. Since most of the data are binned here, the computational time decreases and recognition rate efficiency increases. It can be seen from Figure 8 showing the graph version that there is a linear increase in recognition rate efficiency for different face image factors for Adaptive binning method [ADP].

Table 4 tabulates the results of different imaging factors such as frontal image [Fa], alternative frontal image [Fb], illuminated image [Fc], and aging [Dup1]. The methods compared are Feature extraction method, Eigen Faces method, HGPP and Adaptive binning method. For the last two methods, the comparison is also done by calculating mean values and without calculating mean values. Without mean the computational time increases, but with mean, the checking time reduces and efficiency increases.

	FE Method	**EF Method**	**HGPP Method**		**ADP Method**	
			Mean	**Without Mean**	**Mean**	**Without Mean**
Frontal [Fa]	93.15	95.78	99.12	98.00	95.65	98.45
Alternative Frontal [Fb]	93.25	95.89	98.25	96.13	98.36	97.32
Illumination [Fc]	94.56	95.12	97.58	96.89	98.46	97.48
Aging			91.75	92.00	92.15	93.50
Dup I	85.23	87.21	88.85	89.50	89.1	90.00
Dup II	81.25	82.56	83.65	84.00	84.11	85.15

TABLE 4: Experimental Results

It can be seen From Table 4 for frontal images [Fa] that the recognition rate is good in all the methods listed. But, for the images with background light effects the rate goes down to some extent because of illumination. For an illuminated image, our proposed system gives an increased efficiency than other methods do. Calculation of mean value for generated HGP patterns gives an significant increase in efficiency because of reduced data size in comparison. This can noticed in the figure 9. For aging image factor, when mean is calculated then the efficiency percentage is 95.89% compared with that of 93.85% in HGPP method, when binning concept is used, efficiency percentages remain 97.45 and 95.10 respectively.

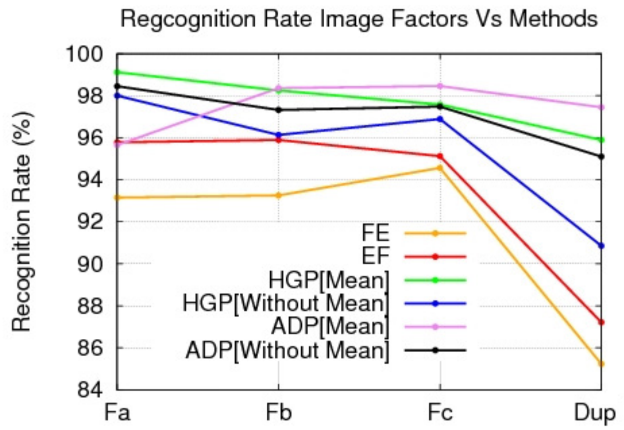


FIGURE 9: Recognition Rate Comparison

6. CONCLUSION

The proposed new system using HGPP and adaptive binning technique gives good recognition rate for different image factors. In this system, the computational complexity arising due to the huge volume of database is reduced and it gives opportunity to extend the database size. A possible future work in this regard could be fixing the appropriate f_{max} value. Further, efficiency could be increased by using modified algorithms for classification and boosting techniques.

7. REFERENCES

1. Baochang Zhang, Shiguang Shan, Xilin Chen & Wen Gao, "Histogram of Gabor Phase Patterns (HGPP). A Novel Object Representation Approach for Face Recognition", IEEE Transactions on Image Processing, vol 16, No.1, pp 57-68, 2007.

2. A.K.Jain, A.Ross and S.Prabhakar, "An introduction to biometric recognition",IEEE TransactioCircuits System. Video Technologies vol 14, no 1 pp. 4-20,2004.

3. Juwei Lu et.al , "Face Recognition Using Kernel Direct Discriminant Analysis Algorithms", IEEE Transcations on Neural Network, vol 14, pp117-126,2003.

4. Tai Sing Lee et. al, "Image Representation Using 2D Gabor Wavelets",IEEE Transactions on Pattern Analysis and Machine Intellignece, Vol 18, pp1-13,1996.

5. Linlin Shel et. al, "A review on Gabor Wavelets for face recognition", Pattern Analysis Application, pp 273-292,2006.

6. Dinu Coltuc,Philippe Bolon and Jean-Marc Chassery , (2006) "Exact Histogram Specification", IEEE Transaction on image processing, vol. 15,No.5,pp.1143-1152.

7. Wenchao Zhang, Shiguang Shan, Xilin Chen, and Wen Gao (2007), "Local Gabor Binary Patterns Based on Mutual Information for Face Recognition", International Journal of Image and Graphics, 7(4) pp: 777-793.

8. Hongming Zhang et. al, "Object detection using spatial histogram features", Image and Vision Computing, pp. 1-15,2006.

9. Sanders J.S. and Fabian A.C. (2001), "Adaptive Binning of X-Ray Galaxy Cluster Images", Institute of Astronomy, Cambridge, pp 1-8.

10. A.Srinivasan, R.S.Bhuvaneswaran, (2008) "Face Recognition System using HGPP and adaptive binning method", Int'l Conf Foundations of Computer Science FCS', pp 80-85.

11. J.S.Sanders, and A.C.Fabian (2001), “Adaptive binning of X-Ray galaxy cluster images”, Monthly Notices of the Royal Astronomical Society Journal, Volume 325, Issue 1, doi: 10.1046/j.1365-8711.2001.04410.x, pages 178–186, July 2001.

12. Hongming Zhang, Wen Gao, Xilin Chen, and Debin Zhao (2006) “Object Detection Using Spatial Histogram Features. Image and Vision Computing”, 24(4) pp: 327-341.

Corner Detection Using Mutual Information

Ibrahim GUELZIM ibr_guelzim@yahoo.fr
Faculty of Sciences, Department of Physics,
GSCM-LRIT Laboratry Associate Unit to CNRST (URAC 29)
Mohamed V-Agdal University
Rabat B.P. 1014, Morocco

Ahmed HAMMOUCH
Faculty of Sciences, Department of Physics,
GSCM-LRIT Laboratry Associate Unit to CNRST (URAC 29)
Mohamed V-Agdal University
Rabat B.P. 1014, Morocco
And GIT-LGE Laboratory, ENSET,
Mohamed V-Souissi University
Rabat, B.P. 6207, Morocco

Driss ABOUTAJDINE
Faculty of Sciences, Department of Physics,
GSCM-LRIT Laboratry Associate Unit to CNRST (URAC 29)
Mohamed V-Agdal University
Rabat B.P. 1014, Morocco

Abstract

This work presents a new method of corner detection based on mutual information and invariant to image rotation. The use of mutual information, which is a universal similarity measure, has the advantage of avoiding the derivation which amplifies the effect of noise at high frequencies. In the context of our work, we use mutual information normalized by entropy. The tests are performed on grayscale images.

Keywords: Computer Vision, Corner Detection, Entropy, Mutual Information, Point of Interest.

1. INTRODUCTION

The detection of points of interest is a fundamental phase in computer vision, because it influences the treatment outcome of several applications: 3D reconstruction, robot navigation, object recognition.

A corner, which is a special case of points of interest, is a point where the direction of a contour changes abruptly. An edge is a transition zone separating two different textures in which the local statistical characteristics of the image may vary slightly [1].

In the literature, several points of interest detectors are proposed. We retain two large families: detectors based on the change of appearance: Moravec [8], SUSAN [2], FAST [3] and detectors based on operators of derivation: Harris [4], Shi and Tomasi [5], Lindeberg [6], Harris-Laplace [7].

The idea of Moravec detector is to determine the average changes in intensity in the neighborhood of each pixel when it moves in four directions: 0°, 45°, 90° and 135°. If there is a significant variation in the average intensity in all directions mentioned, then Moravec decides that the treaty point is a corner [8]. In [2], the authors propose the SUSAN detector where they use a circular mask. The points of the mask that have the same value of intensity of the center, called nucleus, form the USAN area (Univalue Segment Assimilating Nucleus). The information provided by USAN (size, barycenter) allow to detect corners and remove false detections. The final step is the removal of non-maxima. The main advantage of SUSAN is its robustness to noise. In [3], the authors were inspired from SUSAN by using a circular mask, but they only consider the points on the circle. The resulting detector (FAST) has a more isotropic response and is better on repeatability than SUSAN and Harris, but is not robust to noise and depends strongly on the thresholds [3].

Harris and Stephens are based on work of Moravec by considering the Taylor expansion of the intensity function used [4]. The result is a stable detector, invariant to rotation and has good repeatability, by cons it is not invariant to changes of scale and affine transformations and is sensitive to noise. Shi-Tomasi [5] proposed a detector based on principle of Harris detector but by directly computing the minimum of eigenvalues used by Harris.

In [6] Lindeberg proposed a detector invariant to scale changes by a process of convolution with a Gaussian kernel and using the Hessian matrix. In [7] the authors propose an improved Harris detector, invariant to changes of scale and affine transformations. The proposed detector has good repeatability.

In [9] the author proposed an algorithm that detects distinctive keypoints from images and computes a descriptor for them. The interest points extracted are invariant to image scale, rotation. SIFT features are located at maxima and minima of a difference of Gaussians (DoG) function applied in scale space. In [10] the authors propose SURF (Speeded Up Robust Features). It is a scale and rotation invariant detector and descriptor. SURF is based on the Hessian matrix and on sums of 2D Haar wavelet responses and makes an efficient use of integral images [11]. In [12] a revised version of SURF is proposed.

Recently, in [13] the authors propose imbalance oriented selections to detect interest points in weakly textured images. In [14] a new corner detector is proposed based on evolution difference of scale pace. In [15] an algorithm for corner detection based on the structure tensor of planar curve gradients is developed. The proposed detector computes the structure tensor of the gradient and seeks corners at the maxima of its determinant. In [22] the authors use a canny edge detection to present a corner detector based on the growing neural gas. In [23], the authors present a fast sequential method issued from theoretical results of discrete geometry. It relies on the geometrical structure of the studied curve obtained by considering the decomposition of the curve into maximal blurred segments for a given width.

The aim of this paper is to propose a new method of corner detection, invariant to image rotation which is based on a statistical measure that avoids the derivation in order to have better robustness to noise. Our approach is based on the mutual information. Thereafter, we present the proposed method of corner detection with the experimental results and a discussion. Finally we end with a conclusion and prospects.

2. MUTUAL INFORMATION

The mutual information (MI) between two random variables measures the amount of information that knowledge of one variable can make on another. The mutual information between two random variables X = {$x_1, x_2, ..., x_k$} and Y = { $y_1, y_2, ..., y_n$} is:

$$MI(X,Y) = H(X) - H(X/Y) \quad (1)$$

$$= H(Y) - H(Y/X) \quad (2)$$

$$= H(X) + H(Y) - H(X,Y) \quad (3)$$

Such that H is the entropy function and is equal to:

$$H(X) = E[h(x_i)] = -\sum_{i=1}^{k} p_i \log_2(p_i(x_i)) \quad (4)$$

with $p_i = P(X = x_i) / i \in \{1,2,...,k\}$ and $h(x) = -\log(p(x))$.

We have:

$$MI(X,X) = H(X) \quad (5)$$

Mutual information is a positive quantity, symmetric and is cancelled if the random variables are independent.

It follows the principle of no information creation (or Data Processing Theorem):

If g_1 and g_2 are measurable functions then:

$$MI(g_1(X), g_2(Y)) \leq MI(X,Y) \quad (6)$$

The inequality (6) means that no processing on raw data can reveal information.

The MI is a universal similarity measure [16][17][18] which is used in stereo matching [19], image registration[20], parameter selection[21] and other applications.
Figure 1 shows that the mutual Information detects the transition zone separating two different textures in which the local statistical characteristics may vary slightly.

FIGURE 1:a Test image

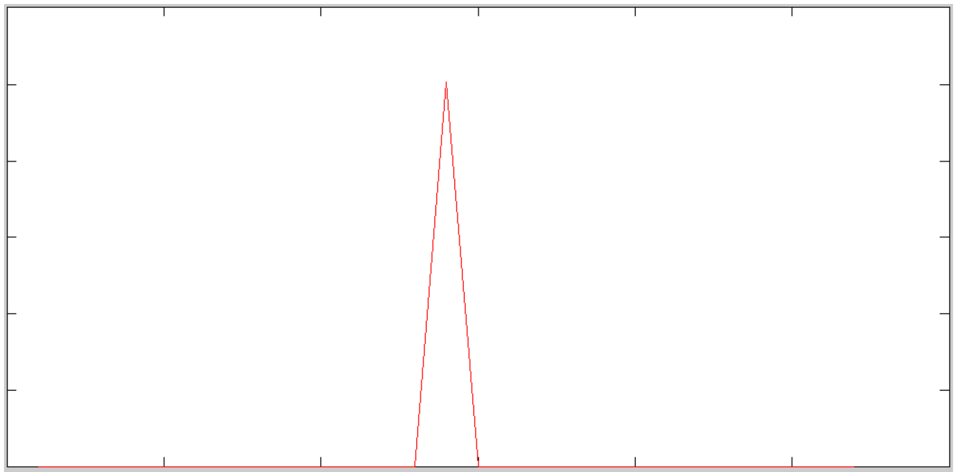


FIGURE 1:B. Values of Mutual information calculated between a window which browses the test image and its right neighbor.

3. PROPOSED METHOD

The proposed method is based on normalized mutual information. It is inspired from Moravec model [8]. Before starting treatment, the first step is the quantification of the values of image pixels in Nl values.
The quantization step is crucial because it allows to homogenize the image areas whose values are relatively close. This is because the calculation of statistics is concerned with the distribution of pixels in the image and not the relative values of the pixels.
The next step is to browse the image by a large window F (Figure 2), and for each pixel in the interior, we calculate the normalized mutual information between its neighborhood W and respective shift of 0 °, 45 °, 90 ° and 135 ° so W1, W2, W3 and W4 (Figure 3).
The mutual information used is normalized by the entropy of the neighborhood W.

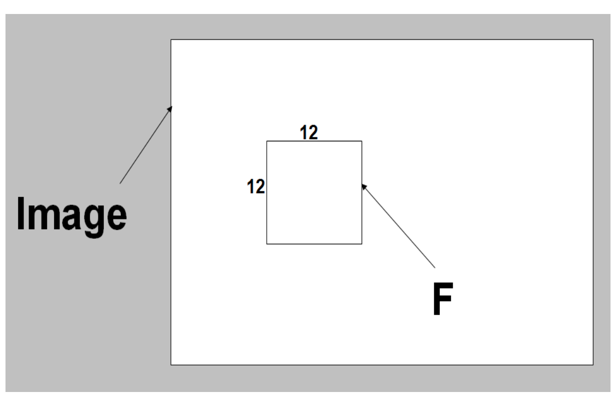


FIGURE 2: Browse the image by a large window F

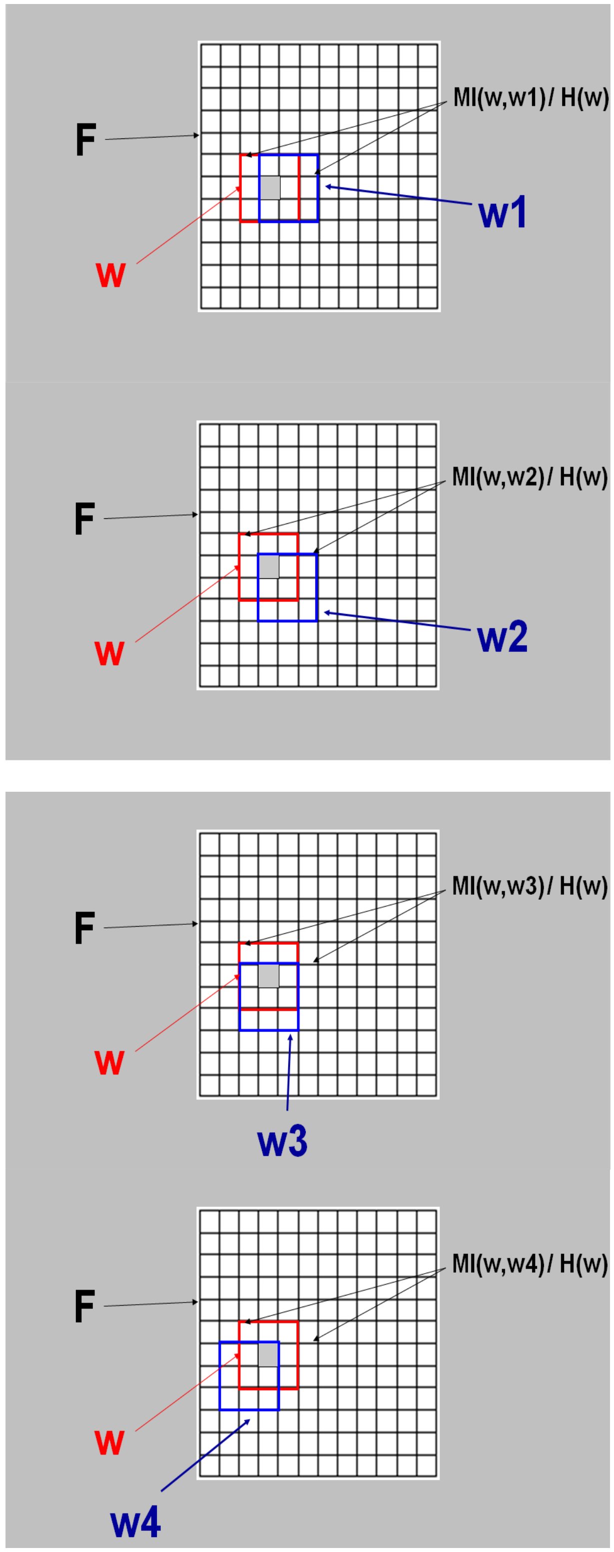


FIGURE 3: Calculation in F of normalized mutual information between w and w1, w2, w3 and w4

If measurements of the normalized mutual information between the neighborhood of the processed pixel and the four other neighborhoods are below a threshold empirically determined, the processed pixel is a corner, otherwise it is not.
Where several corners are detected within the window F, we only keep the point that minimizes the maxima of the normalized mutual information (Figure 4).

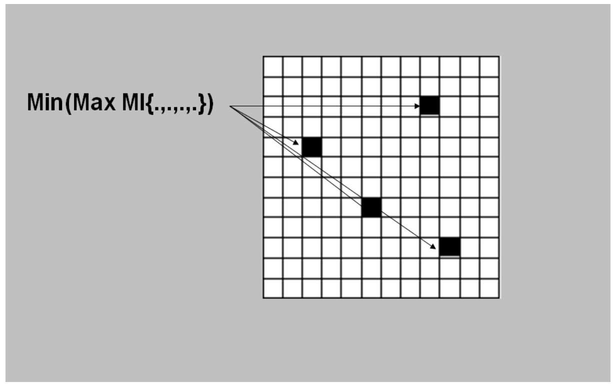


FIGURE 4: Treatment of cases where several corners are detected within the window F.

4. RESULTS AND DISCUSSION
We performed the testing on synthetic and real grayscale images. The results are shown in Figure 5 and Figure 7.
The parameters used: thresholds, window size and number of quantization depend on the image used and are chosen empirically.
In homogeneous areas, the entropy is zero, therefore we can conclude that they cannot contain corners and so we will not divide by zero entropy for normalization.
For areas not homogeneous, if a measure of normalized mutual information between neighbors exceeds the threshold, we can conclude that the processed point is not a corner.
We noisy the normalized images by Gaussian noise amplified by 10, of zero mean and standard deviation equal to 0.2. We notice the good robustness of our method to noise (Figure 6), this is because the proposed method does not use derivation operators that amplifies the effect of noise for high frequencies and are therefore sensitive noise.
In Figure 8, we applied a rotation of 45°, 60° and 90° to the original image. We note the invariance of our method to image rotation. This is due to the statistical tool we use that is invariant to position of pixels but rather to their distribution in neighborhood.

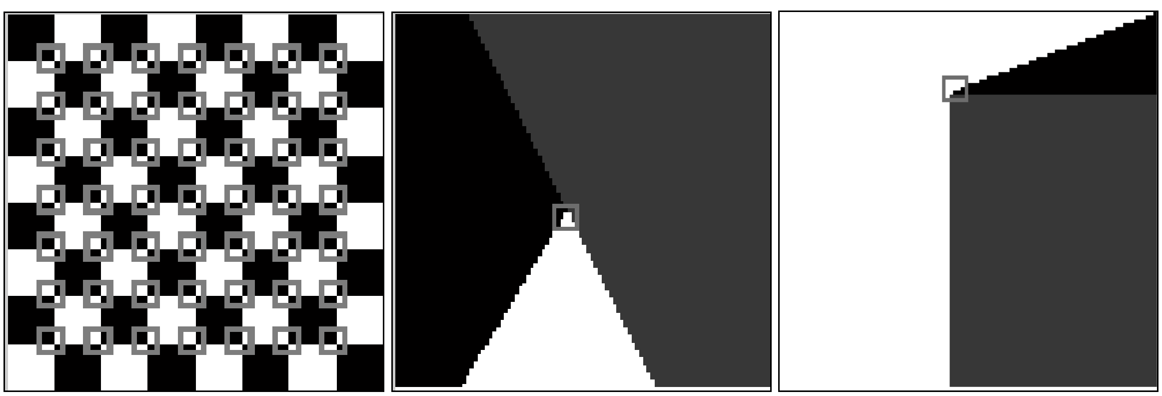


FIGURE 5 : Corner detection results on the original images

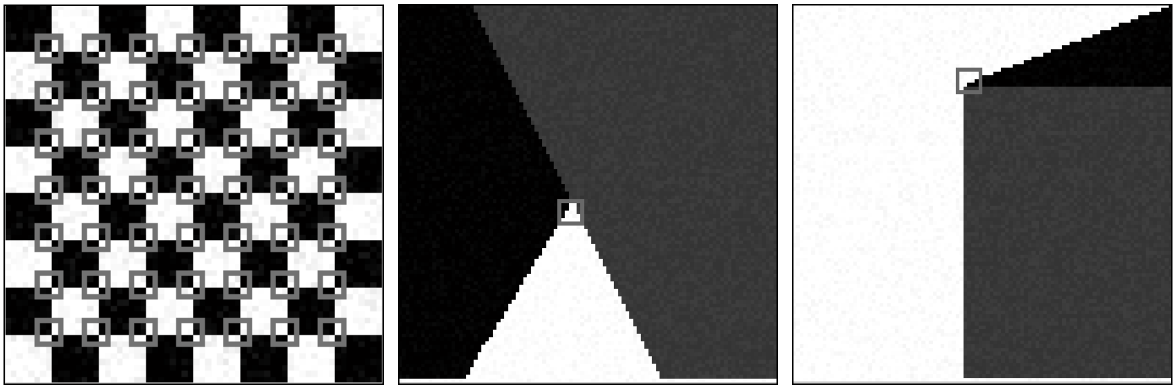
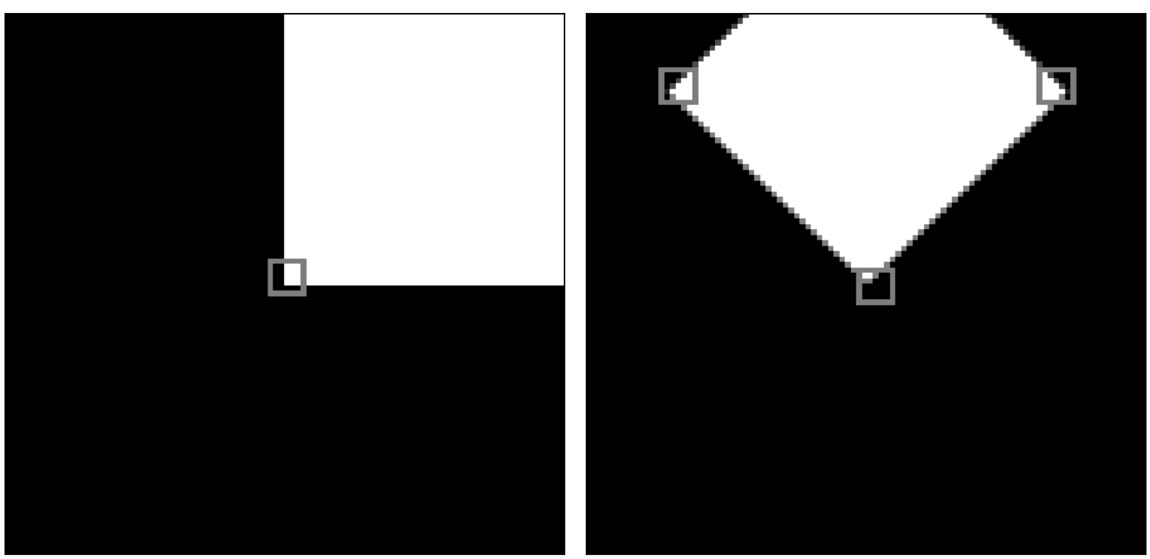
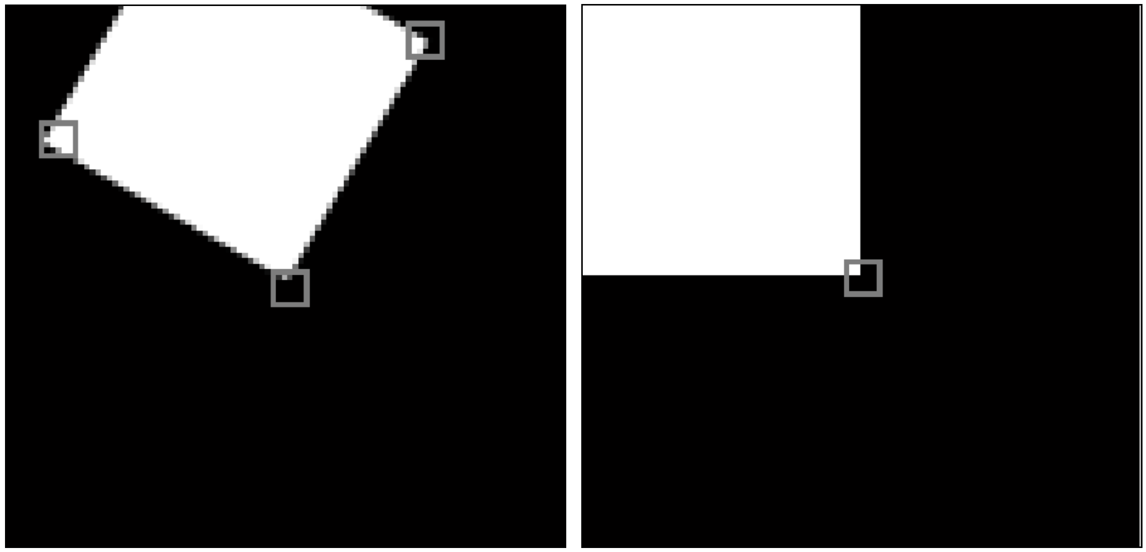


FIGURE 6: Corner detection results on the noised images

FIGURE 7: Examples of corner detection by the proposed method on real images



Original image   Rotation: 45°



Rotation : 60° Rotation : 90°

FIGURE8: Invariance of the proposed method to image rotation

5. CONCLUSION

In this paper we proposed a new corner detection method invariant to image rotation, based on statistical measure which is mutual information. We noted the good performance of our method on synthetic grayscale images. We tested the good robustness of our method to noise. This is explained by the fact that the proposed method does not use derivation operators which are sensitive to noise. The results are promising, which encourages us to apply our method to the mobile robotics knowing that a good detection of points of interest allows good localization of robot and consequently leads to a correct movement.

REFERENCES

1. Keskes, N., Kretz, F., Maitre, H.. Statistical study of edges in TV pictures. IEEE Trans On communications 1979; vol 27. pp:1239-1247.

2. S. M. Smith and J. M. Brady (May 1997). "SUSAN - a new approach to low level image processing.". International Journal of Computer Vision 23. pp:45-78.

3. Edward Rosten (2006), Tom Drummond."Machine Learning for High-Speed Corner Detection". ECCV 2006. pp: 430-443.

4. C. Harris & M. Stephens (1988). A combined corner and edge detector Proceedings of the 4th Alvey Vision Conference: pp:147-151.

5. J. Shi and C. Tomasi. "Good Features to Track". 1994 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94), 1994, pp:593 - 600.

6. T. Lindeberg, "Feature detection with automatic scale selection". International Journal of Computer Vision 30 (2). 1998. pp 77-116.

7. K. Mikolajczyk, C. Schmid. "Scale and affine invariant interest point detectors" Int. J. Comput. Vision 60, Volume 60, Number 1, 2004, pp: 63-86.

8. Moravec H. "Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover", Ph.D. thesis, Stanford University, Stanford, California, May 1980. Available as Stanford AIM-340, CS-80-813 and CMU-RI-TR-3.

9. D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2) (2004). pp:91-110.

10. Bay, H., Tuytelaars, T., Van Gool, L.: "SURF: Speeded Up Robust Features", 9th European Conference on Computer Vision, Graz, Autriche, 7-13 mai 2006. vol. 3954, pp. 404-417.

11. Mozos Ó.M. and Gil A. and Ballesta M. and Reinoso O. “Interest point detectors for visual slam”. In Current Topics in Artificial Intelligence. Vol. 4788. 2007. pp:170-179.

12. Bay H., Andreas Ess, Tuytelaars T. and Van Gool. L, SURF: Speeded Up Robust Features, In Computer Vision and Image Understanding, vol. 110, no 3, 2008, pp. 346-359.

13. Qi Li, Jieping Ye, and Chandra Kambhamettu. Interest point detection using imbalance oriented selection. Pattern Recogn. 41,2. February 2008, pp :672-688.

14. Xiaohong Zhang, Honxing Wang, Mingjian Hong, Ling Xu, Dan Yang, and Brian C. Lovell. Robust image corner detection based on scale evolution difference of planar curves. Pattern Recogn. Lett. 30, 4 (March 2009), 449-455.

15. Xiaohong Zhang, Hongxing Wang, Andrew W. B. Smith, Xu Ling, Brian C. Lovell, and Dan Yang. 2010. Corner detection based on gradient correlation matrices of planar curves. Pattern Recogn. 43, 4 (April 2010), 1207-1223.

16. Skerl D, Tomazevic D, Likar B, Pernus F. Evaluation of similarity measures for reconstruction-based registration in image-guided radiotherapy and surgery. Int J Radiat Oncol Biol Phys. Volume 65, Issue 3 , Pages 943-953, 1 July 2006.

17. Guoyan Zheng and Xuan Zhang. A Unifying MAP-MRF Framework for Deriving New Point Similarity Measures for Intensity-based 2D-3D Registration. IEEE conference. ICPR 2006.

18. D.B. Russakoff, C.Tomasi, T.Rohlfing, Calvin and R.Maurer, Jr. Image Similarity Using Mutual Information of Regions. ECCV 2004.

19. Yong Seok Heo, Kyoung Mu Lee, Sang Uk Lee: Mutual information-based stereo matching combined with SIFT descriptor in log-chromaticity color space. CVPR 2009: 445-452.

20. J. Atif. "Recalage non-rigide multimodal des images radiologiques par information mutuelle quadratique normalisée". Université de Paris XI – Orsay. 29 Octobre 2004.

21. M.A. Kerroum, and A. Hammouch, and D. Aboutajdine. Textural feature selection by joint mutual information based on Gaussian mixture model for multispectral image classification. Pattern Recognition Letters. Volume 31. N°10. pp : 1168-1174. July 2010.

22. W. SUN and X. YANG. Image corner detection using topology learning. The Journal of China Universities of Posts and Telecommunications. Vol. 17 (6): 101-105. December 2010.

23. T. Phuong Nguyen and I. Debled-Rennesson. A discrete geometry approach for dominant point detection. Pattern Recognition. Volume 44, Issue (1). January 2011.

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

- Storage and transmission
- Video coding algorithms and technologies for ATM/p

CALL FOR PAPERS

Volume: 5 - **Issue:** 3 - May 2011

i. Paper Submission: May 31, 2011

ii. Author Notification: July 01, 2011

iii. Issue Publication: July /August 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

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