

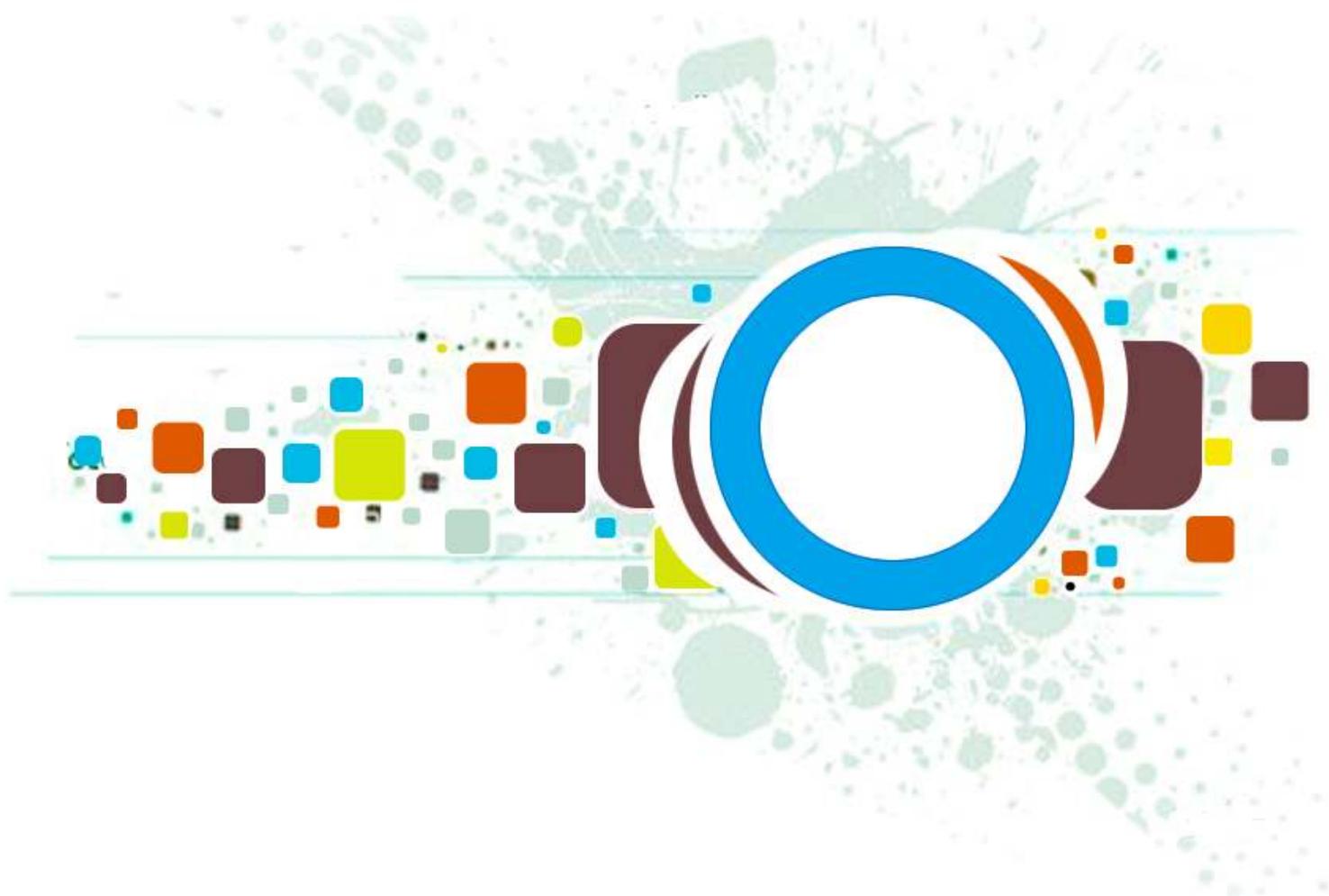
Volume 5 ▪ Issue 3 ▪ August 2011

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 3, 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 3, August 2011

Publishing Date: August 2011

ISSN (Online): 1985-2304

This work is subjected to copyright. All rights are reserved whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication of parts thereof is permitted only under the provision of the copyright law 1965, in its current version, and permission of use must always be obtained from CSC Publishers.

IJIP Journal is a part of CSC Publishers

<http://www.cscjournals.org>

© IJIP Journal

Published in Malaysia

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

CSC Publishers, 2011

EDITORIAL PREFACE

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

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

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

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

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

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

Editorial Board Members

International Journal of Image Processing (IJIP)

EDITORIAL BOARD

EDITOR-in-CHIEF (EiC)

Professor Hu, Yu-Chen
Providence University (Taiwan)

ASSOCIATE EDITORS (AEiCs)

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

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

Assistant Professor Yufang Tracy Bao
Fayetteville State University
United States of America

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

Dr. Huiyu Zhou
Queen's University Belfast
United Kingdom

Professor Yen-Wei Chen
Ritsumeikan University
Japan

Associate Professor Tao Gao
Tianjin University
China

EDITORIAL BOARD MEMBERS (EBMs)

Dr. C. Saravanan
National Institute of Technology, Durgapur West Benga
India

Dr. Ghassan Adnan Hamid Al-Kindi
Sohar University
Oman

Dr. Cho Siu Yeung David
Nanyang Technological University
Singapore

Dr. E. Sreenivasa Reddy

Vasireddy Venkatadri Institute of Technology
India

Dr. Khalid Mohamed Hosny

Zagazig University
Egypt

Dr. Chin-Feng Lee

Chaoyang University of Technology
Taiwan

Professor Santhosh.P.Mathew

Mahatma Gandhi University
India

Dr Hong (Vicky) Zhao

Univ. of Alberta
Canada

Professor Yongping Zhang

Ningbo University of Technology
China

Assistant Professor Humaira Nisar

University Tunku Abdul Rahman
Malaysia

Dr M.MUNIR AHAMED RABBANI

Qassim University
India

TABLE OF CONTENTS

Volume 5, Issue 3, August 2011

Pages

- 236 - 244 Improvement of Objective Image Quality Evaluation Applying Colour Differences in the CIELAB Colour Space
Lisandro Lovisolo, Renata Caminha Coelho de Souza
- 245 - 269 A Survey OF Image Registration
Manjusha Pradeep Deshmukh
- 270 - 282 An Adaptive Two-level Filtering Technique for Noise Lines in Video Images
Baris Baykant Alagoz, Mehmet Emin Tagluk
- 283 - 297 Fast Complex Gabor Wavelet Based Palmprint Authentication
Jyoti Malik, Ratna Dahiya, G Sainarayanan
- 298 - 320 Snow Cover Estimation from Resourcesat-1 AWiFS – Image Processing with an Automated Approach
S SUBRAMANIAM, A V SURESH BABU, E SIVASANKAR, V VENKATESHWAR RAO, G BEHERA
- 321 - 335 An Enhanced Image Warping Technique
Shakeel Doushant Imrith, Maleika Heenaye
- 336 - 351 Comprehensive Performance Comparison of Cosine, Walsh, Haar, Kekre, Sine, Slant and Hartley Transforms for CBIR With Fractional Coefficients of Transformed Image
H.B.Kekre, Sudeep D.Thepade, Akshay Maloo
- 352 - 360 Frequency Domain Blockiness and Blurriness Meter for Image Quality Assessment
Muhammad Tahir QADRI, Tan, K.T, Ghanbari, Muhammad

361 - 370 Particle Swarm Optimization for Nano-Particles Extraction from Supporting Materials
Mohamed abd-ElRahman Abdou

Improvement of Objective Image Quality Evaluation Applying Colour Differences in the CIELAB Colour Space

Lisandro Lovisolo

FEN/DETEL-PEL

Universidade do Estado do Rio de Janeiro

Rio de Janeiro, Brazil

lisandro@uerj.br

Renata C. C. de Souza

FEN/DETEL-PEL

Universidade do Estado do Rio de Janeiro

Rio de Janeiro, Brazil

recaminha@gmail.com

Abstract

Objective full-reference image quality algorithms are often designed considering the luminance component of images. In this work perceived colour distance is employed in a simple and functional way in order to improve these full-reference image quality algorithms. The difference between colours in the CIELAB colour space is employed as perceived colour distance. This quantity is used to process images that are to be feed to full-reference image quality algorithms. This image processing stage consists of identifying the image regions or pixels that are expected to be perceived identically by a human observer in both the reference image and the image having its quality evaluated. In order to verify the validity of the proposal, objective scores are compared with subjective ones for public available image databases. Despite being a very simple strategy, the proposed approach was effective to improve the agreement between subjective and the SSIM (Structural Similarity Index Metric) objective score.

Keywords: Image Quality Evaluation, Color Spaces.

1. INTRODUCTION

Great part of image quality evaluation metrics has been proposed to evaluate achromatic images. This has occurred mainly due to the fact that the spatial resolution of the HVS (Human Visual System) is better for the achromatic (L) channel than for color channels [1]. However, the majority of image applications deal with colorful and not achromatic images. As image capture devices and displays are climbing to a new level (higher luminance, higher contrast ratio, and wider color gamut [2]), the role of color in the ambit of image quality and rendering is growing [2].

The most reliable way of assessing image quality is by means of subjective evaluation, since human observers are the ultimate receivers in most applications [1]. The Mean Opinion Score (MOS) is a reliable and widely used method for subjective quality evaluation obtained from a number of observers [1]. However, for most applications, this method is inconvenient, as it cannot be used in real-time. Due to that, image quality evaluation metrics have been developed. There are different kinds of objective image quality metrics, in this work one considers full-reference metrics, for which both the distorted image (the one to be evaluated) and the original one (reference image) must be available.

Visual image systems work defining color in three dimensional spaces like the RGB, YCbCr, Lab, etc, [3]. While objective metrics are often defined and optimized for the luminance component, there are some previous work that have combined quality scores in the three different color channels of an image for providing one score that is more closely related to subjective ones [4].

A different approach is followed. One focuses on improving the objective evaluation of color images that is provided by traditional full-reference objective metrics that were developed considering just

image luminance. The main idea is to process the image to be evaluated before it is actually fed to the objective evaluation system, aiming at obtaining an objective score that is more close to the subjective one.

The improvement is achieved, by using the concept of Just Noticeable Difference (JND) between colors [3]. The JND concept comes from the definition of the CIELAB color space [3] which is defined with a perceptually uniform measure for color differences in mind [3]. This color space was established by the CIE (*Commission internationale de l'éclairage*) based on MacAdam's Ellipses theory [5]. The area inside each ellipse (defined in the XYZ chromaticity diagram) includes all colors that are visually identical to the color in the ellipse's center. The ellipses radii are known as just noticeable differences (JND) and were translated into the CIELAB color space such that the Euclidean distance between color coordinates provides an approximation to the perceived difference between colors [3].

The colour differences in the CIELAB colour space are employed to compare and process distorted images with original ones before actually feeding the images as input in full-reference objective metric algorithms. The proposed approach is evaluated for the Structural Similarity Index Metric (SSIM) [6], a full-reference (FR) image quality metric.

2. COLOUR DIFFERENCES ANALYSIS AS PREPROCESING FOR IMAGE QUALITY EVALUATION

2.1 CIELAB Color Space

Each one of the three coordinates of the CIELAB color space – L^* , a^* and b^* – represent, respectively the color luminance, the position between red/magenta and green and the position between yellow and blue. To calculate the $L^*a^*b^*$ coordinates from the RGB color space one applies [7]:

$$\begin{aligned} L^* &= 116f(Y/Y_n) \\ a^* &= 500[f(X/X_n) - f(Y/Y_n)], \\ b^* &= 200[f(Y/Y_n) - f(Z/Z_n)] \end{aligned} \tag{1}$$

where

$$f(t) = \begin{cases} t^{1/3}, & \text{for } t > \left(\frac{6}{29}\right)^2 \\ \frac{1}{3}\left(\frac{29}{6}\right)^{2t} + \frac{4}{29}, & \text{otherwise} \end{cases} \tag{2}$$

In (2) X_n, Y_n, Z_n are the CIE XYZ tristimulus values of the white reference and X, Y, Z are related to the RGB color space through the following equation [7]:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \left(\frac{1}{0.17697}\right) \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{3}$$

2.2 Color Difference

The color difference between correspondent pixels in the reference and the distorted images can be defined in the $L^*a^*b^*$ space as [3,5]

$$\Delta E = \sqrt{(L^*_1 - L^*_2)^2 + (a^*_1 - a^*_2)^2 + (b^*_1 - b^*_2)^2}, \quad (4)$$

where (L^*_1, a^*_1, b^*_1) and (L^*_2, a^*_2, b^*_2) are two different colors in the CIELAB color space. L^* , a^* and b^* are the three coordinates of the space and represent, respectively, the color luminance, the position between red/magenta and green, and the position between yellow and blue.

Once one computes these differences between all the pixels in the reference and distorted images one generates a map of differences. In this work it is proposed to use the map of differences to identify the regions where the difference in color is perceptually indistinguishable by human vision, that is, pixels for which their ΔE are below the JND (Just Noticeable Difference).

This map is used to replace the regions in the distorted image that are indistinguishable with respect to the original one by correspondent regions in the original one. This way, a modified, i.e. processed, distorted image is obtained, which is to be fed to the objective metric. The motivation is that if the color difference in a region cannot be perceived, then for the human eye in this region the distorted image is indistinguishable from the reference image and shall not impact the objective metric.

2.3 Applying the Color Difference

Once one identifies the regions where the color difference between the original image and the distorted image are below a given threshold value, the pixels of the distorted image are substituted by the ones in the original image. Define the color of pixel i in the original image as $(L^*_{o,i}, a^*_{o,i}, b^*_{o,i})$ and in the distorted image as $(L^*_{d,i}, a^*_{d,i}, b^*_{d,i})$. Then, for a given pixel i

$$\text{If } \Delta E_i < \text{threshold then } (L^*_{d,i}, a^*_{d,i}, b^*_{d,i}) \leftarrow (L^*_{o,i}, a^*_{o,i}, b^*_{o,i}). \quad (5)$$

This process is illustrated in Figure 1. After this color-space processing stage, the full-reference objective metric is computed considering the original and the modified distorted image.

Similar ideas for quality assessment of colour images or videos exist in the literature. In [8] the inclusion of psycho visual models after processing and transforming the images is employed. In [9] some effects of the human visual system are applied independently on the reference and distorted image previous to image comparisons. In [4] quality scores for the different colour channels of an image are combined into a single score. Our work is much simpler and differs from these as it tries to include the influence of color on image quality in the pixel domain. It is not intended as an image-quality metric but it aims at improving the results obtained with other metrics.

3. EXPERIMENTS

One aims at evaluating if the application of the JND concept (if considering that the difference in colors is not noticeable, as the processing stage described in the previous section) is able to improve the similarity between objective and subjective scores for the distorted images. For that purpose it is investigated if the objectives scores obtained for the processed images have a larger correlation with and a smaller difference to the subjective scores than the ones obtained for the original distorted images, i.e., without the proposed processing stage. This strategy follows the recommendations for the evaluation of objective quality metrics [10].

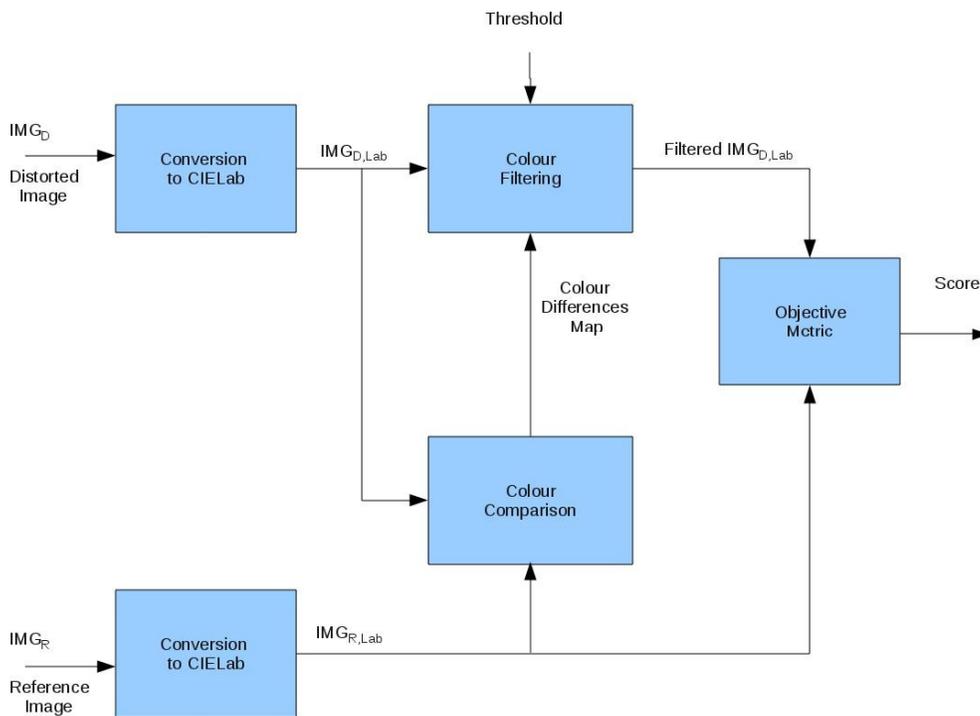


FIGURE 1 – Image processing for improving its quality analysis by full-reference image-quality metrics.

3.1 Materials Employed

As exposed above, in order to evaluate an objective quality metric, a subjective database is welcomed. In subjective image quality databases, distorted images are presented to human subjects that grade the quality of the images accordingly to some prescribed criteria. For the results presented here, two databases were used: the LIVE subjective database [11] containing 808 images (29 reference images and 779 distorted images) and the correspondent DMOS (Difference Mean Opinion Score), and IVC database [12] containing 120 distorted images from 10 reference images and their MOS.

3.2 Objective Metric

To evaluate the proposed image processing stage it is employed as objective metric the SSIM [6]. The SSIM is designed assuming that people evaluate impairments between images by perceiving their structural information such as mean, variance, and covariance of intensity values in regions of the reference and the distorted images. As it is depicted in Figure 1, in the proposed approach, the calculation of SSIM considers the channel L^* (luminance in the $L^*a^*b^*$ space). As the JND does not assume an exact value for all applications and viewing conditions [13], experiments were held in a set of threshold values ranging from 0 (meaning results without using the proposed approach) to 7 in steps of 0.2.

3.3 Methodology Employed for Evaluation of the Results

The problem one must deal now is how to compare the results obtained by two different objective metrics. Since visual human system inspired metrics as the SSIM try to mimic the subjective evaluation, one should evaluate how close the objective scores with respect to subjective one are. For that purpose, two “metrics of similarity” between the objective and subjective scores were employed. The Pearson Correlation between objective and subjective scores and the RMSE (Root Mean Squared Error) between the same are calculated to evaluate the performance of the proposal. A nonlinear mapping of the subjective scores to the objective notes is made prior to

these calculations, following the methodology used by the VQEG (Video Quality Experts Group) [10].

The Pearson correlation is given by

$$R = \frac{\sum_{n=1}^N (subj_n - \overline{subj})(obj_n - \overline{obj})}{\sqrt{\sum_{n=1}^N (subj_n - \overline{subj})^2} * \sqrt{\sum_{n=1}^N (obj_n - \overline{obj})^2}} . \quad (6)$$

Where, $subj_n$ and obj_n are respectively the subjective and objective scores of a given image n in a set of N images and \overline{subj} and \overline{obj} are the mean of the subjective and objective scores for the N images. The Pearson correlation is commonly used to evaluate how close the objective metric is to the subjective score [10].

The RMSE is computed using

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (subj_n - obj_n)^2} . \quad (7)$$

The *RMSE* analyzes the precision of the objective metric. While the correlation can be employed to compare the behavior of one metric with respect to the other, one employs the *RMSE* to evaluate how far the results provided by the objective metric are to the image quality scores obtained in subjective tests.

4. RESULTS

The correlation and *RMSE* between SSIM and the subjective scores for different thresholds used in the colors comparisons for both LIVE and IVC databases are shown respectively in Figures 2 and 3 in black continuous line. The upper graphs in those figures show the correlation between the objective and subjective qualities indexes for threshold values ranging between 0 and 7. The graphs in the bottom present the *RMSE* between these scores. Note that a threshold equal to zero actually represents that the proposed image processing approach is not used prior to objective quality evaluation. In order to evaluate the consistence of the proposed approach, the databases were divided into subsets of images. These subsets divide the IVC database into four groups and the LIVE database into five groups in a random manner. This was applied three times, so that 12 subsets for IVC database and 15 subsets for LIVE database were defined. For these subsets correlation and *RMSE* were also calculated (using equations (6) and (7), respectively) and are shown in Figure 2 and 3 in dashed blue lines.

One shall highlight the shapes of the curves, which follow a consistent pattern, showing hills in the correlation graphs and valleys in the *RMSE* graphs for intermediate threshold values. Although these ranges are different for the two databases, this difference is expected due to the different conditions in which the subjective scores were obtained, the observed patterns attest the effectiveness of the proposed approach for improving color image objective evaluation.

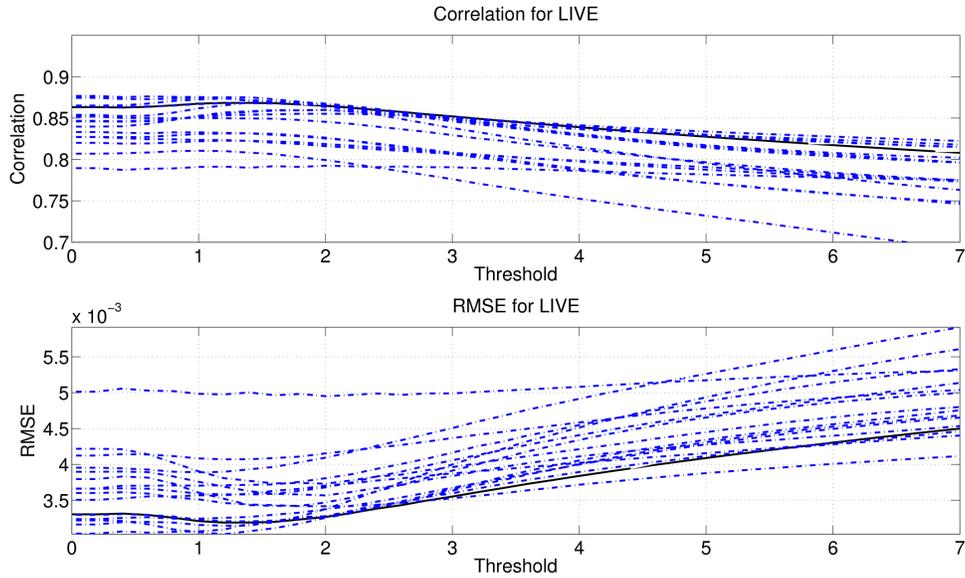


FIGURE 2 – Results of objective image quality obtained with the proposed approach for LIVE Database.

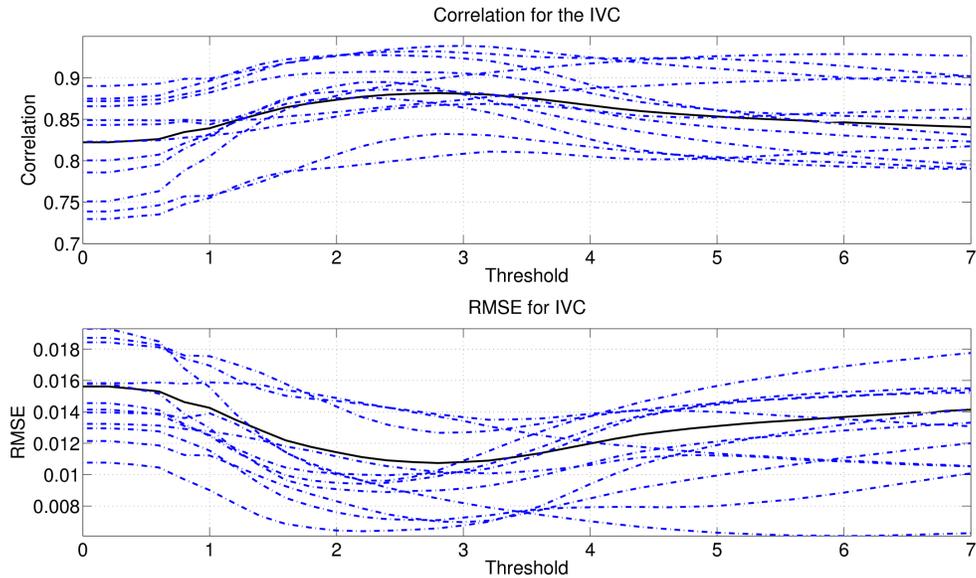


FIGURE 3 – Results of objective image quality obtained with the proposed approach for IVC Database.

Table 1 shows a synthesis of the results obtained. In Table 1, are presented the improvements obtained using the proposed approach with respect to not using it. It shall be highlighted that the improvements in the SSIM for the LIVE database are smaller than the ones obtained for the IVC. However, SSIM and LIVE are both originated from the same research group and the constants used in the SSIM were tuned using the LIVE database. However, one should observe that regardless the database, and regardless if one considers the whole database or a subset of it, one observes that adjusting the value of the threshold considered to compare color differences affords a gain in the correlation between SSIM and the subjective metric. In addition, the RMSE is minimized at the same threshold value.

IVC	Threshold	0	2,8
	Correlation	0,82	0,88
	Improvement regarding JND=0	-	7,22%
LIVE	Threshold	0	1,2
	Correlation	0,86	0,87
	Improvement regarding JND=0	-	0,60%

TABLE 1: Comparison between correlations using JND for SSIM.

5. CONCLUSION

In this work, we used the difference between colors computed in the CIELAB color space in the context of objective image quality evaluation. It was proposed to use this difference for identifying the regions where both the original and distorted images are indistinguishable. Once this is done, we substitute these regions in the distorted images by the correspondent ones in the original one before feeding the images to a full-reference metric algorithm. The proposed method is simple, and only requires some additional preprocessing before the image pair (original and distorted) is sent to the calculation of the objective metric. Experiments using different subjective image quality public available databases have demonstrated that this approach improves the results for the SSIM full-reference image quality evaluation metric.

6. REFERENCES

- [1] Z. Wang, H. R. Sheikh and A. C. Bovik, *The Handbook of Video Databases: Design and Applications*, 2nd ed., Boca Raton, FL, USA. CRC Press, 2003.
- [2] T. Fujine, T. Kanda, Y. Yoshida, Y., M. Sugino, M. Teragawa, Y. Yamamoto, and N. Ohta, N., "Bit-Depth Needed for High Image Quality TV-Evaluation Using Color Distribution Index," *Journal of Display Technology*, vol. 4, no. 3, pp. 340–347, Sep. 2008.
- [3] G. Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. 2nd. ed. John Wiley & Sons, New York, USA, 1982.
- [4] A Toet and M.P. Lucassen, "A New Universal Color Image Fidelity Metric", *Displays*, vol 24, no 4-5, pp 197-207. 2003.
- [5] D. L. MacAdam, "Specification of small chromaticity differences in daylight," *Journal of the Optical Society of America*, vol. 33, no. 1, Jan. 1943.
- [6] Z. Wang and A.C. Bovik ., "Image Quality Assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [7] J. Schanda, *Colorimetry: Understanding the CIE System*, Wiley-Interscience, 2007.
- [8] J. Lubin, "A visual discrimination model for imaging system design and evaluation", in E. Peli (ed.), *Vision models for target detection and recognition*, World Scientific Publishing, Singapore, 1995.
- [9] S. Daly. *The visible difference predictor: An algorithm for the assessment of image fidelity*, *Proc. SPIE Conference on Human Vision and Electronic Imaging XII*, p. 2. 1992.
- [10] VQEG. *Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II*. [Online] <http://www.vqeg.org/>, Aug. 2003.

- [11] H. R. Sheikh, et al. *LIVE Image Quality Assessment Database Release 2*, <http://live.ece.utexas.edu/research/quality>, 2009.
- [12] P. Le Callet and F. Autrusseau. *Subjective Quality Assessment IRCCyN/IVC Database*, <http://www.irccyn.ec-nantes.fr/ivcdb/>, 2005.
- [13] M. Melgosa, M. M. Pérez, A. Yebra, R. Huertas and E. Hita, "Some reflections and recent international recommendations on color-difference evaluation", *Óptica Pura y Aplicada*, vol. 34, 2001.

A SURVEY OF IMAGE REGISTRATION

Manjusha Deshmukh

*Asst. Prof., Saraswati College of Engineering
Kharghar Navi Mumbai, 410210, India*

manju0810@yahoo.com

Udhav Bhosle

*Prof., Rajiv Gandhi College of Engineering,
Andheri, Mumbai, India*

udhavbhosle@gmail.com

Abstract

In image processing, for example, when combining the information content of image, we are interested in the relationship between two or more images. The analysis of this relationship usually becomes tractable once a correspondence is set up between the images. Image registration is the task of setting up this correspondence. This paper overviews the theoretical aspects of an image registration problem. The purpose of this paper is to present a survey of image registration techniques. Registration is a fundamental task in image processing used to match two or more pictures taken, for example, at different times, from different sensors, or from different viewpoints. It geometrically aligns two images the reference and sensed images. Specific examples of systems where image registration is a significant component include matching a target with a real-time image of a scene. Various applications of image registration are target recognition, monitoring global land usage using satellite images, matching stereo images to recover shape for navigation, and aligning images from different medical modalities for diagnosis.

Keywords: Image Registration, Image Transformation, Image Mosaicing.

1. INTRODUCTION

Image registration is establishment of correspondence between images of the same scene. Many image processing applications like remote sensing for change detection, estimation of wind speed and direction for weather forecasting, fusion of medical images like PET-MRI, CT-PET etc need image registration. Image registration is a process of aligning two images acquired by same/different sensors, at different times or from different viewpoint. To register images, we need to determine geometric transformation that aligns images with respect to the reference image. The most common transformations are rigid, affine, projective, perspective and global. A variety of approaches have been published on this topic. Over the years, a large range of techniques has been developed for various types of problems. All these techniques have been independently studied for different applications. This paper organizes this research by establishing the relationship between the variations in the images and the type of registration techniques which can most appropriately be applied. All registration techniques explained in paper are useful for understanding the merits and relationships between the wide variety of existing techniques and for supporting in the selection of the most appropriate technique for a particular problem.

A widespread survey of image registration methods was published in 1992 by Brown [1]. A comprehensive survey of image registration methods is presented by Barbara Zitova and Jan Flusser [2]. They have classified the image registration techniques as area based methods and feature based methods. J. B. Antoine Maintz presented a survey of medical image registration in 1998.[3].Leszek Chmielewski presented various methods of image registration in 2001.[4].Mohammad Essadiki presented a technique for combining panchromatic and multispectral spot images [5].Subunku in his thesis presented various entropy based image registration techniques[6].S. K. Bose presented various tools for medical image registration [7].J.

Flusser used moment based approach to correct affine distortion, he has also done degraded image analysis to locate invariants in images[8,9]. Sangit Mitra and B.S. Manjunath explained various contour based approaches for multispectral image registration in their different papers[11]. Cahill, N D Williams, C.M. Shoupu propose an approach to incorporate spatial information into the estimate of entropy to improve multimodal image registration [15]. J.P.W. Plum presented another survey on medical image registration [16]. Frederik Maes and Andre Collignon apply mutual information to measure the statistical dependence or information redundancy between the image intensities of corresponding voxels in both images [17]. Xiaoxiang Wang and Jie Tian in their paper proposed a mutual information based registration method using gradient information rather than pixel intensity information [18]. Frederik Maes, Andre present novel histogram based method for estimating and maximizing mutual information between two multimodal and possibly multiband signals [19]. J. P. Queiroz developed method for automatic registration of satellite images acquired on different dates, for both geometric and radiometric correction with respect to reference image [20]. Shannon's paper, titled "A mathematical theory of communication" is widely accepted as the origin of information theory. In this paper, Shannon used probability theory to modal information sources, i.e. data produced by a source is treated as a random variable [21]. Haim Schweitzer in his paper proposed that large collection of images can be indexed by projections on a few "eigenfeatures", the dominant eigenvectors of the images covariance matrix [22]. Ma Debao and Liwagao introduced the new matrix characteristic methods like eigenvalues and eigenvectors and achievable accuracy is derived theoretically and verified by tests using simulated interferometric data.[23]. Haim Schweitzer demonstrated that eigenspace based algorithm registers multiple images and produces improved eigenfeatures[24]. Wen Cao and Bicheng proposed PCAT (principal component analysis transform) and WPT (wavelet packet transform) for remotely sensed image fusion[25]. Yu-Te Wu, Takeo Kanade, Ching-Chung Li and Jeffrey Cohn proposed that their wavelet-based algorithm produced better motion estimates with error distributions having a smaller mean and smaller standard deviation [26]. Tarek A El-hazawi explained wavelet based image registration [27]. Hala S. Own and Aboul Ella Hassanien in their paper presents an efficient image registration technique using the Q-shift complex wavelet transform (Q-shift CWT). The experimental results proved that the proposed algorithm improves the computational efficiency and yields robust and consistent image registration compared with the classical wavelet transform [28]. Azhar Quddus and Otman Basir proposed a novel, fully automatic, multistage wavelet-based image registration technique for image retrieval applications. They used multiscale wavelet representation with mutual information (MI) to facilitate matching of important anatomical structures at multiple resolutions. The proposed approach has several novel aspects including the use of MI in multistage wavelet domain [29]. Stone, Harold S., Le Moigne, Jacqueline; McGuire and Morgan proposed fast image registration by progressively registering wavelet representations at different resolutions[30]. Ghazaw presented wavelet based image registration on parallel computing[31]. Jiangsheng explained image matching techniques by using Radon transform [32].

2. TRANSFORMATIONS

Image registration consists of establishing correspondence means matching of identical shapes in related image pair. This requires geometric transformation of one image into another. Change in viewpoint or relative motion between the camera and object planes introduces distortion in the features of an image e.g. a circle may appear ellipse when observed from non-fronto-parallel vantage point. However, certain features of object shape remain intact even after such transformations. These features are called invariants. The fundamental characteristic of any image registration technique is the type of spatial transformations or mapping used to properly overlay two images. The most common transformations are rigid, affine, projective, perspective, and global [1, 2].

2.1 Euclidean or Rigid Body Transformation

$$p' = Rp + t$$

$$\text{where } R = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

is a rotation matrix and $t=[t_x \ t_y]^T$ is translation vector. p' and p are the transformed and original 2-D points, respectively, represented in non homogeneous coordinates $[x' \ y']^T$ and $[x \ y]^T$ respectively. Euclidean invariants are lengths and angles.

2.2 Similarity Transform

$$p' = sRp + t$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = s \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ 0 \end{bmatrix}$$

Here s is a scaling factor. Similarity invariants are angles, ratios of lengths, and ratios of areas.

2.3 Affine Transformation

The most commonly used registration transformation is the affine transformation which is sufficient to match two images of a scene taken from the same viewing angle but from different position. It is composed of scaling, translation, and rotation. It is global transformation which is rigid. Affine transformations are more general than rigid.

$$p' = Ap + t$$

The general 2D affine transformation

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} t_x \\ t_y \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Angles and lengths are not preserved. Parallel lines remain parallel.

2.4 Projective Transform

This is the most general geometric transformation. Here, two 2-D points p' and p (represented in homogeneous coordinates), are related by a 3×3 non-singular transformation matrix (homography).

$$p' = Hp$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Projective invariants include the cross ratio of four collinear points, or four concurrent lines [1, 2, 3].

3. REGISTRATION ALGORITHMS

Image registration algorithms can be classified in various ways, like based on modality, intensity or methods used for registration. Barbara Zitova and Jan Flusser [2] classified the image registration techniques as area based methods and feature based methods. Area based methods are preferably applied when in images prominent details are absent and distinctive information is provided by gray levels / colors rather by local shapes and structure. Feature based matching methods are applied when local structural information carried by image intensities are more. These methods make use of image features derived by feature extraction algorithm. Point of sharp variations such as edges, corners, contours, surfaces, point of intersection etc. what carries valuable information about images are used for matching.

3.1 Pixel Based Method

Cross-correlation is the basic statistical approach of registration. It is often used for template matching or pattern recognition in which the location and orientation of a template or pattern is found in picture. Cross correlation is a similarity measure or match metric. For template T and image I, where T is small compared to I, the two dimensional normalized cross-correlation function measures the similarity for each translation[1,2,3,4].

$$C(u, v) = \frac{\sum_x \sum_y T(x, y) I(x - u, y - v)}{\sqrt{[\sum_x \sum_y I^2(x - u, y - v)]}}$$

If template matches the image, then cross correlation will have it's peak at C(i, j). Cross correlation must be normalized since local image intensity would otherwise influence the measure.

Implementation on Dataset-1



a	b	c
---	---	---

FIGURE 1: (a) Original image SCOE-1. (b) translated image SCOE-2. (c) Correlation plot for SCOE 1 & SCOE 2.

Observations

Two main drawbacks of the correlation-like methods are the flatness of the similarity measure maxima (due to the self-similarity of the images) and high computational complexity. The maximum can be sharpened by preprocessing or by using the edge or vector correlation. It is used in next section.

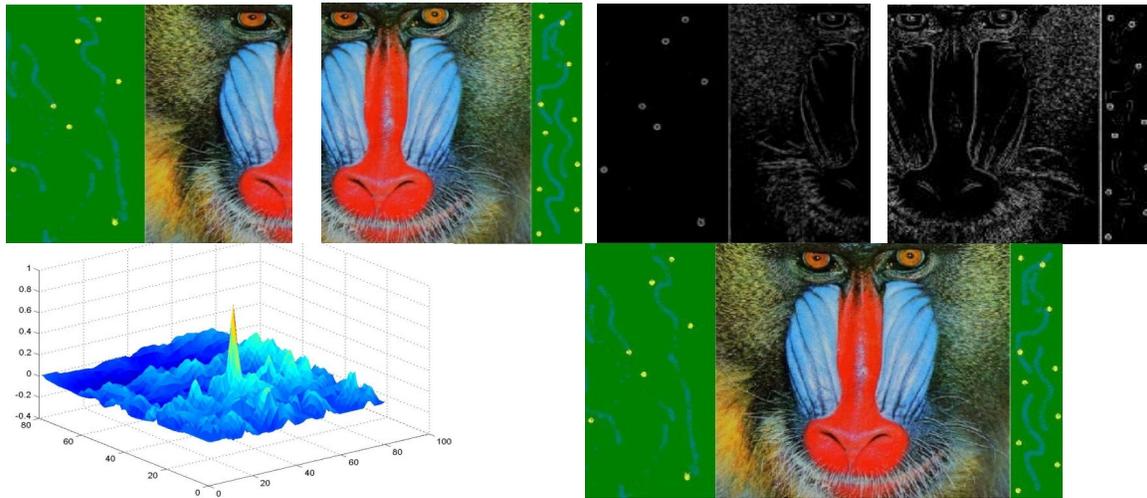
3.2 Feature Based Methods or Point Mapping Method

Feature based matching techniques do not use the grey values to describe matching entities. It makes use of image features derived by feature extraction algorithm. The purpose of feature extraction is to abstract substantial information from original data input and filter out the redundant information. Features are selected which are likely to be uniquely found in both images and more tolerant of local distortions. Computing of proper transformation depends on these features. Therefore sufficient number of features must be detected to perform calculation. After detecting features in each image, they must be matched. This technique is primary approach to register two images whose type of misalignment is unknown. This occurs if class of transformations cannot be easily categorized as translations or rigid-body movements. In this we can use landmarks and match them using general transformation. The method of point mapping consist of three stages-

- Computing features in the images
- Control points in reference image are corresponded with feature points in data image.
- Spatial mapping.

Control points for point matching play an important role in this method. Control points may be corners, line of intersections, points of locally maximum curvature on contour lines, centers of windows having locally maximum curvature and centers of gravity of closed-boundary regions [1, 2, 14]. Under this method we have computed edges of images in color vector space by using color gradients For a scalar function $f(x, y)$, the gradient is a vector pointing in the direction of maximum rate of change of f at coordinates (x, y) .

Implementation on Dataset 1



a	b	c	d
e	f		

FIGURE 2: (a) mar1 image adopted from MATLAB toolbox. (256 X 256). (b) mar2. (c) Features extracted from MAR-1.(d) Features extracted from MAR-2.(e) Correlation plot between (c) & (d). (f) registered mar1 and mar2

Implementation on Dataset-2

In this experiment image correction of rotated and stretched image of clown is done by point mapping method. In this method features are extracted using color gradients. Invariants are identified and projective transformation is applied on image to correct it.



a	b
c	

FIGURE 3: (a) Clown 1 original image(adopted from MATLAB toolbox).(b) : Clown 2 original image stretched and rotated.(c) Image corrected by projective transformation.

Observation

Feature based registration methods overcomes the limitation of correlation method and are often in use, particularly because of their easy hardware implementation, which makes them useful for real-time applications. Disadvantage of the these methods refers to the 'remarkableness' of the window content. There is high probability that a window containing a smooth area without any prominent details will be matched incorrectly with other smooth areas in the reference image due to its non-saliency. The features for registration should be preferably detected in distinctive parts

of the image. Windows, whose selection is often not based on their content evaluation, may not have this property.

3.3 Registration Based on High Level Features (Contour Based Image Registration)

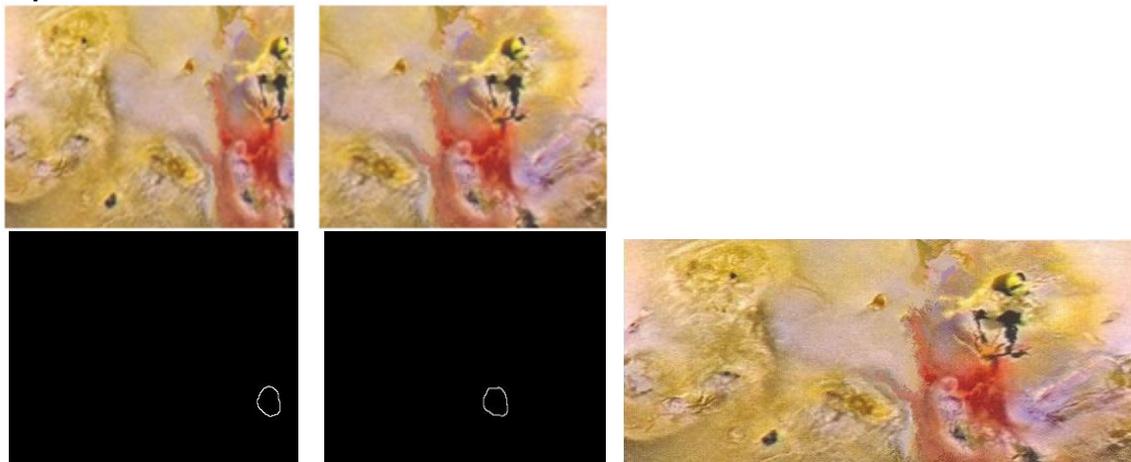
This method makes use of high statistical features for matching image feature points. For extracting regions of interest from the image color image segmentation is used. Given a set of color of interest, obtain mean of color 'm'. For segmentation classify each RGB pixel in image as having a color in the specified range or not. Euclidean distance is used to measure similarity. 'z' is an arbitrary point in RGB space and T is threshold. Euclidean distance between 'z' and 'm' is given by

$$D(z, m) = \|z - m\| = [(z - m)^T (z - m)]^{1/2}$$

$$= [(z_R - m_R)^2 + (z_G - m_G)^2 + (z_B - m_B)^2]^{1/2}$$

Here RGB denotes RGB components. The locus of points such that $D(z, m) \leq T$ is a solid sphere of radius T and point contained within, or on the surface of the sphere satisfy specified color criterion. Coding these two sets of points in the image with black and white produces a binary, segmented image. After segmentation remove noise by 'Gaussian' filter. Threshold blurred image and then obtain contour of image. For registration of jup1 and jup2 contours of red colors are extracted by above said method. Centroid of contours i.e. (x, y) coordinates of centre of gravity of jup2 are corresponded with centroid of jup1 contour. Similar procedure is repeated to extract contours of berry images. Angle between x-axis and major axis of contour is calculated to find angle of rotation.

Implementation on Dataset-1



a	b	
c	d	e

FIGURE 4 : (a) original image jup-1 (256 X256).(b) translated image jup-2(256 X256). (c) Counter extracted from jup-1(d) Counter extracted from jup-2(e) Mosaic of jup-1 and jup-2

Implementation on Dataset-2

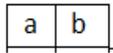
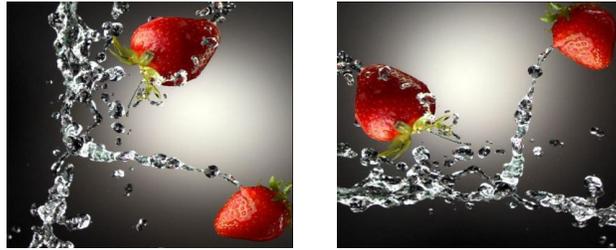


FIGURE 5: (a) Original image Berry1(256 X256). (b) Berry2 rotated by 90 deg. (256 X256).

Actual angle of rotation = 90 deg, Calculated angle of rotation = 88.85 deg.

Observations:

Feature based methods do not use the gray values for matching and hence overcomes the limitations of spatial methods. Feature based method filter out the redundant information. Accuracy of this method is more but the limitation is , it is manual and slow.

3.4 Multimodal Image Registration Using Mutual Information

Multispectral image registration is also called as multimodal analysis. Images of the same scene are acquired by different sensors. The aim is to integrate the information obtained from different source streams to gain more complex and detailed scene representation. Different types of application are, in remote sensing fusion of information from sensors with different characteristics like panchromatic images, offering better spatial resolution, color/multispectral images with better spectral resolution, or radar images independent of cloud cover and solar illumination. Medical imaging applications are, combination of sensors recording the anatomical body structure like MRI, ultrasound or CT with sensors monitoring functional and metabolic body activities like PET, SPECT. Result can be applied, in radiotherapy and nuclear medicine.Registration of multispectral / multisensor images is a challenging area. In general such images have different gray level characteristics and simple techniques such as those based on area correlation cannot be applied directly. This section is an attempt to solve this difficult problem by employing a basic concept from information theory.

Entropy of image

For a high dimensional discrete random variable $X = (X_1, X_2, \dots, X_d)$ that has a probability mass function of $P(x_1, x_2, \dots, x_d)$, the entropy is

$$H(X) = \sum_{x_1, x_2, \dots, x_d} P(x_1, x_2, \dots, x_d) \log \frac{1}{P(x_1, x_2, \dots, x_d)}$$

From the properties of entropy, entropy doesn't depend on value of random variable, but only depends on distribution. This is explained with following example.

E.g. consider image shown in following figure 6.

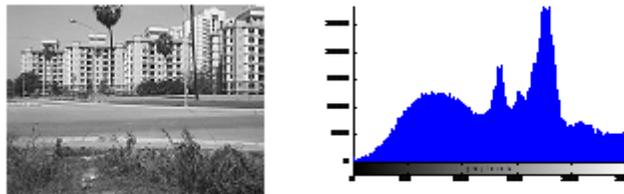
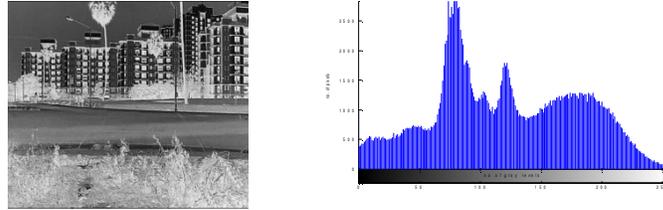


FIGURE 6: (a)The image of size 500 X 500 with 256 gray levels. (b)Histogram of image.

Entropy (gray image) = 7.7540

By applying bijective mapping to the intensity values, we get synthetic image as shown below in figure 7.



a	b
---	---

FIGURE 7: (a) Synthetic image of size 500X500 (left) with 256 gray levels.(b)Histogram of image

Synthetic Entropy (synthetic image) =7.7549

From Figure 6 and 7, it is easy to see that entropy of image does not change even if histogram has changed. Even after randomly shuffling pixels of image, entropy of image remains same. From above discussion it is clear that natural images contain less uncertainty. In this type of images pixel intensity values depend on neighboring pixels. In other words, in natural image the value of pixel is likely to be close to some of its neighbors. Hence this dependency reduces the total entropy [6].

Entropy as Alignment Measure

The objective of using entropy as alignment measure is simple. Corresponding features extracted from the images should become statistically more dependent with better alignment. This is explained with following example.



a	b	c	d
e	f	g	h

FIGURE 8: (a) Subject. (b) subject rotated by 5°. (c) Subject rotated by 20°. (d) Subject with rotated neck. (e) Joint histogram a Vs a. (f) Joint histogram a Vs b. (g) Joint histogram a Vs c. (h) Joint histogram a Vs d .

In Figure 8, scatter plots (joint histogram) display pixel intensity value pairs from all images. Notice that, since in Figure e. images are aligned, pixel samples cluster around x-y line. At bad alignment, samples are scattered, i.e. joint histogram is more dispersed. This indicates that if

alignment is good joint distribution tends to be sharper with peak at good alignment. This illustrates misregistration measured by dispersion of 2-D histogram of image intensities[6].

Mutual Information to Describe Dispersive Behavior of 2-D Histogram

Mutual information measures statistical dependence between two random variables. Mutual Information criteria presented here states that, mutual Information of image intensity values of corresponding voxel pairs is maximum if images are geometrically aligned. Let A and B represent random variables and $P_A(a)$ and $P_B(b)$ represents its marginal probability distributions. Let $P_{AB}(a,b)$ represents joint probability distribution

then A & B are independent if $P_{AB}(a,b) = P_A(a) * P_B(b)$ Mutual Information I(A, B) is given by

$$I(A, B) = \sum_{a,b} P_{AB}(a,b) \log [P_{AB}(a,b) / \{ P_A(a) \cdot P_B(b) \}]$$

Mutual Information is related to entropy by following equations.

$$I(A, B) = H(A) + H(B) - H(A, B)$$

Where $H(A, B)$ is joint entropy and $H(A/B)$ & $H(B/A)$ are conditional entropy

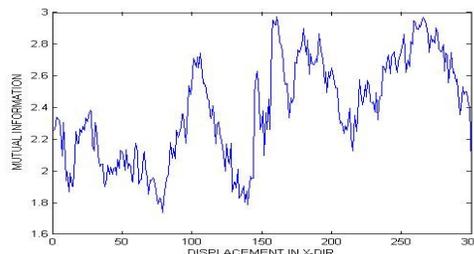
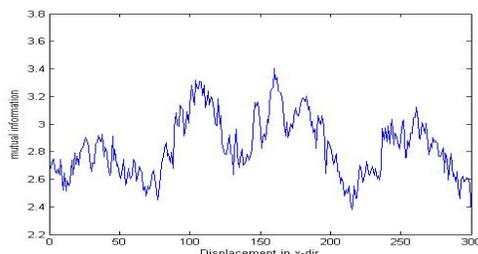
$$H(A) = - \sum_a P_A(a) \cdot \log P_A(a)$$

$$H(A/B) = - \sum_{a,b} P_{AB}(a,b) \cdot \log P_{AB}(a,b)$$

$$H(A/B) = - \sum_{a,b} P_{AB}(a,b) \cdot \log P_{A/B}(a/b)$$

Consider that x & y is image intensity values of a pair of corresponding voxel in two images which are to be registered. Intensities x and y are related through the geometric transformation T_α defined by registration parameter α . The MI registration criteria states that the images are geometrically aligned by the transformation T_α for which $I(A,B)$ is maximum. This is explained in following example. Figure 9 (d) and (e), shows 2-D histogram of image intensity values in non-registered and registered position. We observe that at good alignment i.e. at $x=160$, joint distribution is sharper and shows peak. Mutual Information between SCOE 1 & SCOE 2 is 3.4026 and matching point is at 160. Similarly Mutual Information between SCOE 1 & SCOE 3 is 2.97 and matching point is at 161.

Implementation on Dataset-1

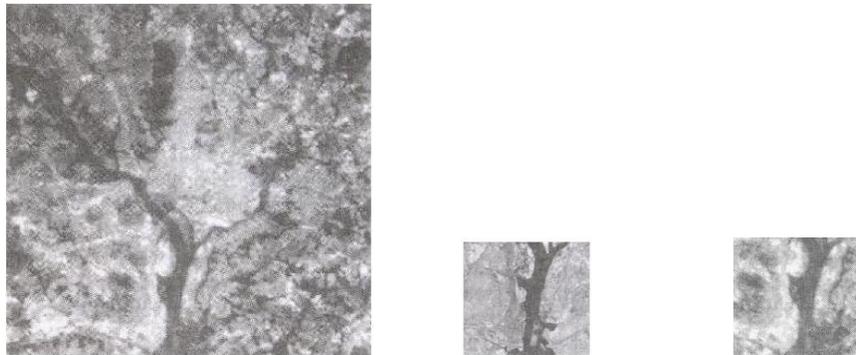




a	b	c
d	e	
f	g	

FIGURE 9 : (a-c)Sample images SCOE-1, SCOE-2 and SCOE-3.(d) I (A, B) as a function of mis registration for SCOE 1 and SCOE 2.(e) I (A, B) as a function of misregistration for SCOE 1 and SCOE 3.(f) mosaic of SCOE1 and SCOE2.(g) mosaic of SCOE1 and SCOE3 in multiple spectrums

Implementation on Dataset-2

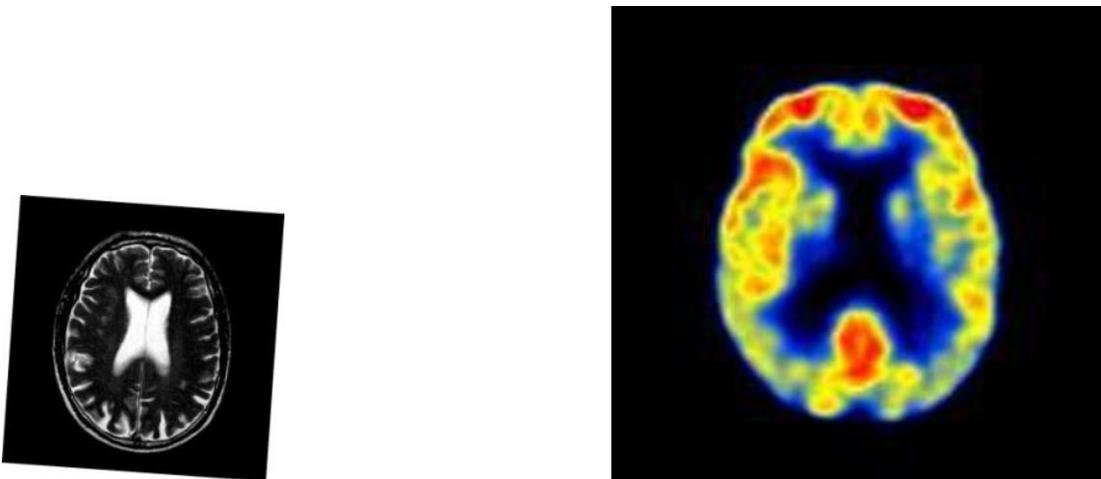


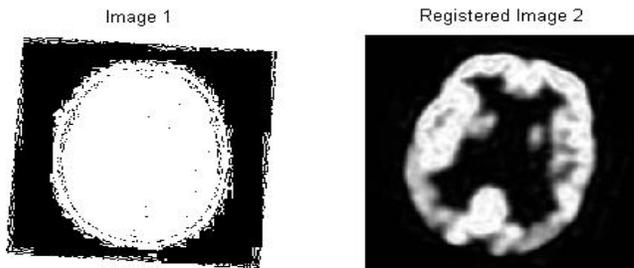
a	b	c
---	---	---

FIGURE 10: (a) Image in thermal infrared bands (450 X 450). (b)Template in near infrared (150 X 150).(c)Matched template (150 X 150).

For images a, b and c, Template matched at $x = 124$ and $Y = 300$
 Mutual information = 0.8522

Implementation on Dataset 3 to Correct Rotation and Scale

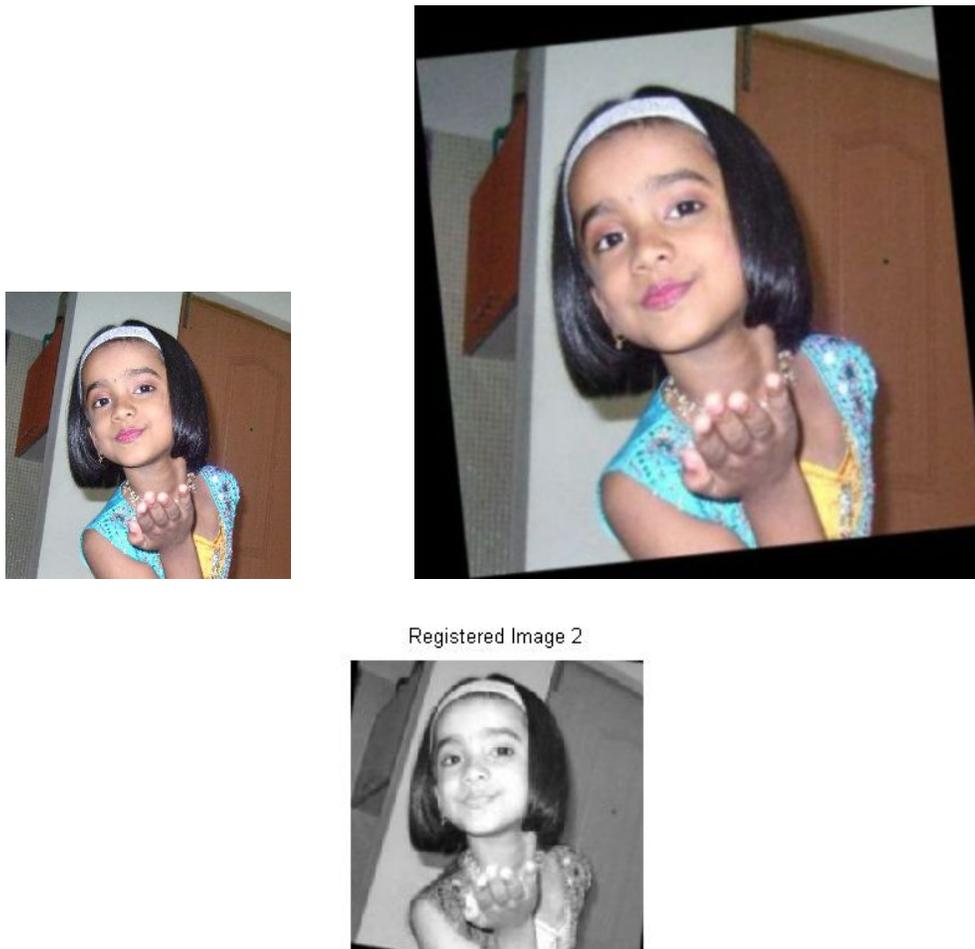




a	b
c	d

FIGURE 11: (a) MRI image (230 X 230) rotated 4° .(b) CT image (512 X 512).(c-d)Registered CT image(233 X 233)

Implementation on Dataset 4 to Correct Rotation and Scale



a	b
c	

FIGURE 12: (a) Prutha image (230 X 230) .(b) Prutha image (512 X 512) rotated by 6° .(d)Registered Prutha image(231 X 231)

Observation

It is observed that proposed method yields a more accurate registration than any other registration method. But this method has its own limitations. When images are of low resolution, when images contain little information, or when the region of overlap is small then mutual information result in mis-registration .It has one more limitation of speed , when time is an important constraint one cannot use this method . Although it has some limitations entropy and mutual information are best approaches for multimodal image registration.

3.5 Image Registration in Frequency Domain

Correlation theorem has one useful property. Correlation theorem states that, the Fourier transform of the correlation of two images is the product of Fourier transform of one image and complex conjugate of Fourier transform of other. The Fourier transform of an image $f(x,y)$ is a complex function, each function value has real part $R(\omega_x,\omega_y)$ and an imaginary part $I(\omega_x,\omega_y)$ at each frequency (ω_x,ω_y) of frequency spectrum.

$$F(\omega_x, \omega_y) = |F(\omega_x, \omega_y)| e^{-j\varphi(\omega_x, \omega_y)}$$

Where $|F(\omega_x, \omega_y)|$ is magnitude , and $\varphi(\omega_x, \omega_y)$ is phase angle

$$|F(\omega_x, \omega_y)|^2 = R^2(\omega_x, \omega_y) + I^2(\omega_x, \omega_y)$$

$$\varphi(\omega_x, \omega_y) = \tan^{-1} \left[\frac{I(\omega_x, \omega_y)}{R(\omega_x, \omega_y)} \right]$$

Cross power spectrum of two images is defined as

$$F(\varphi_x, \varphi_y) = \frac{F1(\varphi_x, \varphi_y) F2^*(\varphi_x, \varphi_y)}{|F1(\varphi_x, \varphi_y) F2^*(\varphi_x, \varphi_y)|}$$

Shift theorem guarantees that phase of cross power spectrum is equivalent to the phase difference between the images. If we represent the phase of cross power spectrum in it's spatial form, i.e. by taking the inverse Fourier transform of the representation in the frequency domain, then we will have a function which is an impulse, that is approximately zero everywhere except at displacement which is needed to optimally register two images. Above method is used to register images having only translation.

Implementation on Dataset-1 to Correct Translation



FIGURE 13: (a) Edges extracted from SCOE-1. (b) Edges extracted from SCOE-2. (c) cross power spectrum for SCOE 1 & SCOE 2 in frequency domain.

Observation

It is observed that in frequency based method accuracy is more than correlation method but less as compared to other methods. But if we extract image features and then apply Fourier method

accuracy increases. In frequency domain it should be noted that some form of interpolation must be used[1,2].

3.6 Wavelet Transform Based Image Registration

Registration of image using wavelets is being discussed in this section. Image Registration using wavelets has already been studied by many researchers. In this step feature that is extracted using wavelet or other transforms used as input and correspondence between these features will be recognized. To find out correspondence between two images so that transform model can be calculated using these correspondence points is very important part of overall problem. Cross correlation or Mutual Information could be used as a measure of similarity

Wavelet transform decomposes an image into various sub images based on local frequency content. Using discrete wavelet transform (DWT), a function $f(t)$ can be represented by

$$f(t) = \sum_{j,k} a_{j,k} \psi_{jk}(t)$$

Where $a_{j,k}$ are wavelet coefficients, $\psi_{j,k}(t)$ are basis function, j is scale, k is translation of mother wavelet $\psi(t)$. Two dimensional DWT can be obtained by applying DWT across rows and columns of an image. The two dimensional DWT of image $f(x,y)$ is

$$f(x, y) = \sum_{j,k} C_{J_0}(k, l) \phi_{j,k,l}(x, y) + \sum_{S=H,V,D} \sum_{J=J_0}^{\infty} \sum_{k,l} D_j^S[k, l] \psi_{j,k,l}^S(x, y)$$

Where C_{J_0} is approximation coefficient, $\phi_{j,k,l}(x,y)$ is scaling function, D_j^S is set of detail coefficients and $\psi_{j,k,l}$ is set of wavelet function. The DWT coefficients are computed by using a series of low pass filter $h[k]$, high pass filters $g[k]$ and down samplers across both rows and columns. The results are the wavelet coefficient the next scale. The filter bank approach to calculate two dimensional dyadic DWT is shown in Figure 14. The wavelet coefficients are of smaller spatial resolution as they go from finer scale to coarser scale. The coefficients are called the approximation (A), horizontal detail (H), vertical detail (V) and diagonal detail (D) coefficient.

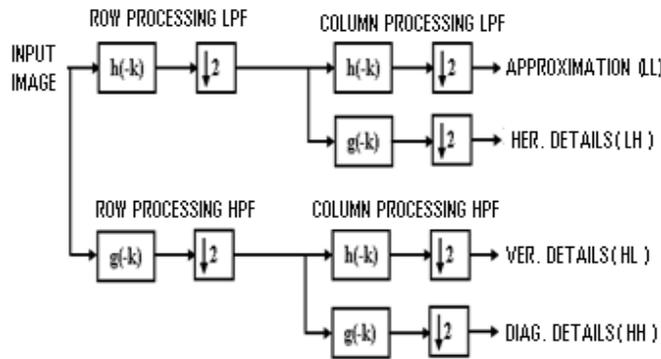


FIGURE 14: Two-dimensional orthogonal wavelet decomposition

In wavelet transformation due to sampling, the image size is halved in both spatial directions at each level of decomposition process thus leading to a multi-resolution signal representation. The decomposition and reconstruction of wavelet pyramid of source images are based on Mallat's theories[25,26,27,28,29,30,31].



a	b
c	d

FIGURE 15: (a)Original image SCOE-1(256X256).(b)Translated image SCOE-2(256X256).(c)Wavelet decomposition at level 2 by Harr wavelet for SCOE-1.(d) Wavelet decomposition at level 2 by Harr wavelet for SCOE-2.

Mutual Information as a Measure of Similarity

Origin of MI is from information theory, MI is an entropy-based concept and denotes the amount of information that one variable can offer to the other. Mutual information measures statistical dependence between two random variables. Mutual Information criteria presented here states that, mutual Information of image intensity values of corresponding voxel pairs is maximum if images are geometrically aligned. Let A and B represent random variables and $P_A(a)$ and $P_B(b)$ represents its marginal probability distributions. Let $P_{AB}(a, b)$ represents joint probability distribution then A & B are independent if $P_{AB}(a, b) = P_A(a) * P_B(b)$

Mutual Information I(A, B) is given by

$$I(A, B) = \sum_{a,b} P_{AB}(a,b) \log [P_{AB}(a,b) / \{ P_A(a).P_B(b) \}]$$

Mutual Information is related to entropy by following equations.

$$I(A, B) = H(A) + H(B) - H(A, B)$$

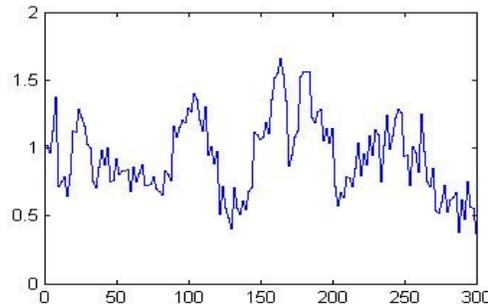


FIGURE16: I (A, B) as a function of mis registration for SCOE 1 and SCOE 2

Observations

It is observed that combinational approach of wavelet and mutual information gives better results as compared to wavelet - correlation combination and Fourier based image registration. Even

wavelet mutual information combination can be used in case of multimodal image registration. If we involve mutual information as similarity criteria, speed decreases and hence we can not use this combination in applications where speed is required.

3.7 Hotelling Transform Based Image Registration

In this section automatic method of template matching using Hotelling transform as well as Hotelling transform for image alignment is presented. The proposed method is validated on pair of remotely sensed scenes and medical images.

Principal Component Analysis

In this section we present Hotelling Transform (Principal Component Analysis) for template matching and image alignment to register images. The proposed method is validated on pair of remotely sensed scenes and medical images. Proposed method for template matching is compared with other algorithms.

Principal component analysis is one of the most frequently used dimension reduction method. Principal component analysis also called as Hotelling Transform. Hotelling Transform is based on stastical properties of vector representations. The material discussed here can be used as the basis for describing sets of images, that are registered spatially, but their corresponding pixel values are different. If we have 'n' component images having different pixel values, These images can be treated as a unit by expressing each group of 'n' corresponding pixels as a vector. Let x_1, x_2, \dots, x_n are values of first pixel in each of the n' images then 'n' elements can be expressed as follow

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

Where, x is column vector. This one vector represents one common pixel in all two image. If images are of size P X Q, there will be total of N=P*Q such n- dimensional vectors.

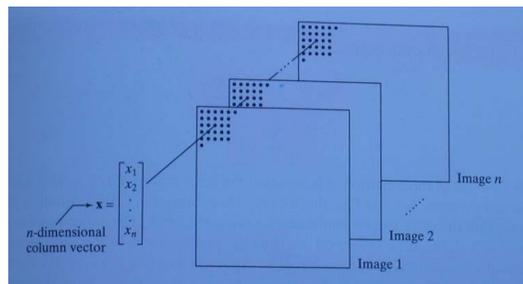


FIGURE 17: Forming a vector from corresponding pixels in a stack of images of the same size

We can assume these vectors as random quantities,
Mean vector of population is

$$m_x = E[x]$$

The covariance matrix of vector population is

$$C_x = E\{(x - m_x)(x - m_x)^T\}$$

Element C_{ij} of C_x is the variance of x_i , the i^{th} component of x vectors and element C_{ij} is covariance between components x_i and x_j . The matrix C_x is symmetric and real. For a sample of N vectors from a random population, the mean vector and covariance matrix can be given by expression

$$m_x = \frac{1}{N} \sum_{k=1}^N x_k$$

Thus covariance matrix can be estimated as

$$C_x = \frac{1}{N} \sum_{k=1}^N x_k x_k^T - m_x m_x^T$$

Since C_x is real and symmetric, it is possible to find a set of N ortho-normal eigenvectors. Let e_i and λ_i be eigenvectors and corresponding eigenvalues of C_x where $i = 1, 2, \dots, N$. 'A' is a matrix whose rows are eigenvectors of covariance matrix C_x . Then A is ordered so that the first row of A is eigenvectors corresponding to the largest eigenvalue, and last row corresponds its smallest eigen value. If we use A as transformation matrix to map the x 's into y . Then y is given by

$$Y = A(x - m_x)$$

Above expression of 'y' is called Hotelling Transform or Principal Component Transform. Alternatively, the Hotelling Transform can be viewed as the discrete version of the Karhunen-Loeve transform (KLT). This transform has some useful properties. Mean of 'y' vector is zero i.e.

$$m_y = E(y) = 0$$

Covariance matrix of y 's is estimated as

$$C_y = A C_x A^T$$

It is observed that C_y is a diagonal matrix. Elements of C_y along main diagonal are eigenvalues of C_x .

$$C_y = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \dots & \lambda_2 & \dots \\ \dots & \dots & \dots \\ \dots & \dots & \dots & \lambda_n \end{bmatrix}$$

λ_j are eigenvalues of C_x . The important property of Hotelling Transform is we can reconstruct vector x from vector y .

$$\hat{x} = A^T y + m_x$$

The error between vector x and reconstructed vector \hat{x} can be minimized by using the eigenvectors corresponding to the largest eigenvalues. The fact that zero values are everywhere except along the main diagonal in C_y indicates that the components of the transformed vector y are uncorrelated i.e. the correlation previously existing between the different components of random vector x has been removed in transformed domain. Therefore, if the input is split into blocks and Hotelling Transform is applied block wise, the coding may be more efficient since the data in the transformed block are uncorrelated.

Geometrical Interpretation of Hotelling Transform (PCA)

Here we are giving a geometrical interpretation of transform coding. For this, we use 2-D vectors instead on N -D vectors. Hotelling Transform can also be used to align region or boundaries with the eigenvectors of the object. We have to form two-dimensional vectors from the coordinates of the boundary or region. Figure 18(a) shows that vectors are formed from the coordinates of the pixel in the object or we can use coordinates of points on the boundary. The resulting vectors are treated as 2-D population of random vectors. Let each pixel in the object is a 2-D vector $x = (a, b)^T$, where a and b are coordinate values of that pixel with respect to x_1 and x_2 axis.

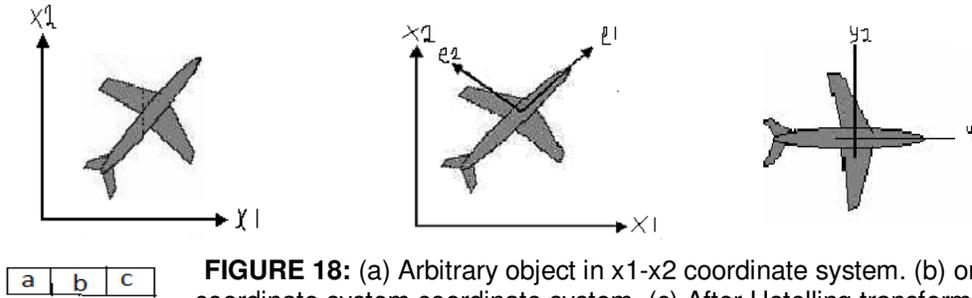


FIGURE 18: (a) Arbitrary object in x_1 - x_2 coordinate system. (b) origin of new coordinate system coordinate system. (c) After Hotelling transform, the object is aligned with its principle axes.

These vectors are used to calculate mean vector m_x and covariance matrix C_x . Then the effect of using equation of Hotelling Transform is to establish a new coordinate system. The origin of this new coordinate system is at the centroid of the population and its axes are in the direction of eigenvectors of C_x . It is shown in figure 18 (b). It is observed that transformation in equation $y=A(x-m_x)$ is a rotation and translation that aligns the object with eigenvectors. Two eigenvectors are perpendicular. The 'y' axis is also called as eigen axis. The ability of Hotelling Transform to align the object with its principal axis provides a reliable means for removing the effects of rotation. This is a rugged alignment procedure which uses all coordinates of object (region or boundary) to compute the transformation matrix and aligns the data in the direction of its principal spread [15].

Implementation on Dataset-1 for Image Alignment

In this experiment, Hotelling Transform is applied to binary images. Fig. 19 (a) shows gradient image of MRI of head and fig 19 (b) shows same image but head is at translated and rotated position.

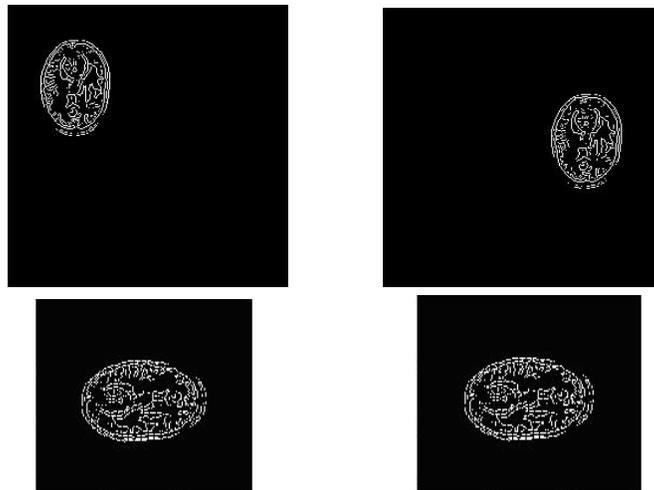


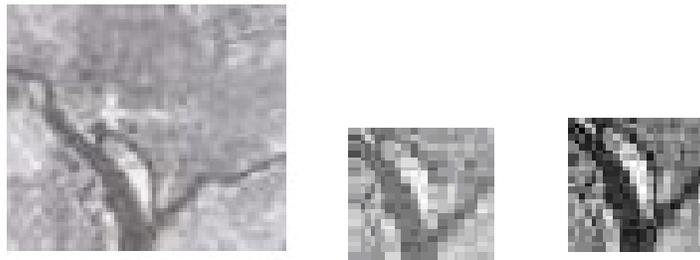
FIGURE 19: (a) MRI image of brain (400X400). (b) Translated and rotated version of same image (400X400). (c) Hotelling transform of the image a. (d) Hotelling Transform of the images b.

Figure 19(c,d) shows Hotelling Transform of MRI image and its rotated version. It can be seen that Hotelling Transform automatically positioned the two images in such a way that corresponding points are practically aligned. This removes rotational effects. Translational effects are also removed since the object is centered on its mean. After this alignment images can be fused to get complete information for diagnosis. In this section we have given background theory

of principal component analysis and discussed implementation details. It is observed that if we use eigenvalues and eigenvectors for registration of images accuracy is more precise than that of other methods. We can use his method for template matching. It is observed that Hotelling Transform can be used to align region or boundaries with the eigenvectors of object but this method can be applied only to regions and boundaries. In terms of future work, effort could be directed at the problems of an unknown image scaling and unknown global rotation between the target and reference images. Both are outstanding issues requiring a solution and should prove useful for multimodal images. There are possible solutions to the problem that could be investigated in future work.

Implementation on Dataset -2 for Template Matching

Template matching is the process of finding the location of sub image, called a template, inside an image. Template matching involves comparing a given template with windows of the same size in an image and identifying the window that is most similar to the template. The accuracy of a template matching process depends on the accuracy of metric used to determine the similarity between a template and a window. There are many similarity measures known to produce best results .Hotelling Transform can be used for template matching. It is observed that when template match with the window of the same size in an image then eigenvector corresponding to that location shows largest eigenvalue and eigenvalue of second eigenvector shows minimum value.



a	b	c
---	---	---

FIGURE 20: (a) Sample image (50 X 50). (b) Template (25 X 25). (c) Matched template (25 X 25).

4. COMPARATIVE STUDY OF IMAGE REGISTRATION METHODS.

All methods explained in the paper are implemented on the three datasets . This made comparison of all the methods easy.

Implementation on Dataset-1 on Dataset 1

Main application of image registration is in remote sensing for image mosaicking of surveyed area for monitoring of global land usage; landscape planning etc. A camera typically has a limited field of view. A lens with a wide field of view incurs substantial distortion. In addition, capturing the entire scene with the limited camera resolution compromises the image quality. Image mosaicking algorithms register or stitch a sequence of images into a composite image.



FIGURE 21: Sample images and mosaic of VR images from H1 to H 16

Observations

Location of maximum match point is shown in table

TABLE I: Location of Maximum Match using Different Methods

S.N.	Image combination	Correlation Method	Fourier Transform Method	Mutual Information Method	Principal Component Analysis	Wavelet Transform Method	Feature Based Method
1	H1-H2	63	62	65	62	59	61
2	H2-H3	110	111	113	112	113	114
3	H3-H4	62	63	57	62	63	63
4	H4-H5	79	77	78	80	80	82
5	H5-H6	147	147	150	147	148	150
6	H6-H7	93	92	89	93	93	95
7	H7-H8	72	74	73	74	74	75
8	H8-H9	73	74	73	74	74	74
9	H9-H10	88	88	90	89	88	90
10	H10-H11	83	84	84	83	83	82
11	H11-H12	79	79	82	80	80	80
12	H12-H13	81	81	78	81	81	82
13	H13-H14	85	84	89	89	83	89
14	H14-H15	128	126	129	128	125	129

TABLE II: Location of Maximum Match using Different Methods

S.N.	Image combination	Correlation Method	Fourier Transform Method	Mutual Information Method	Principal Component Analysis	Wavelet Transform Method	Feature Based Method
1	H1-H3	169	167	171	169	145	168
2	H2-H4	171	172	167	170	165	170
3	H3-H5	136	133	140	136	133	135
4	H4-H6	223	233	235	223	224	224
5	H5-H7	231	237	240	232	233	232
6	H6-H8	161	158	150	160	162	162
7	H7-H9	141	140	130	140	141	140
8	H8-H10	158	156	160	159	155	157
9	H9-H11	143	143	130	142	142	142
10	H10-H12	163	162	150	164	163	164
11	H11-H13	160	161	132	159	158	159
12	H12-H14	165	165	165	165	163	164
13	H13-H15	210	201	210	210	210	209

Implentation on Dataset-2

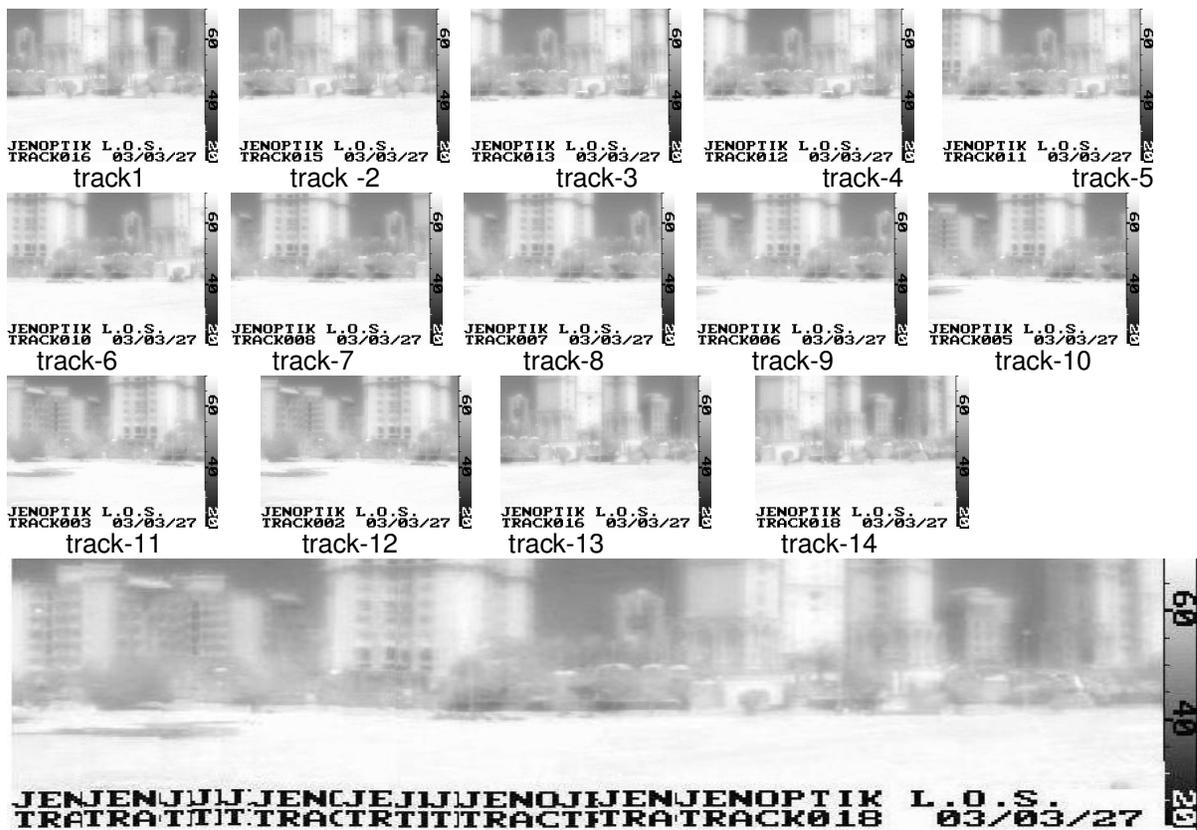


FIGURE 22: Sample images in infrared band and mosaic of IR images from track-1 to track-14

Observation

Location of maximum match point is shown in table.

TABLE III : Location of Maximum Match using Different Methods

S.N.	Image combination	Correlation Method	Fourier Transform Method	Mutual Information Method	Principal Component Analysis	Wavelet Transform Method	Feature Based Method
1	T1-T2	44	45	46	45	46	44
2	T2-T3	51	50	48	49	50	50
3	T3-T4	19	21	22	20	20	20
4	T4-T5	22	21	21	20	19	20
5	T5-T6	20	20	21	20	22	20
6	T6-T7	57	59	50	57	58	58
7	T7-T8	30	32	32	30	34	31
8	T8-T9	22	24	20	24	22	23
9	T9-T10	20	20	11	19	22	18
10	T10-T11	60	63	70	61	60	62
11	T11-T12	28	28	30	27	28	27
12	T12-T13	50	50	51	51	50	49
13	T13-T14	27	27	28	27	28	27

TABLE IV: Location of Maximum Match using Different Methods

S.N.	Image combination	Correlation Method	Fourier Transform Method	Mutual Information Method	Principal Component Analysis	Wavelet Transform Method	Feature Based Method
1	T1-T3	94	95	99	95	96	94
2	T2-T4	71	71	71	71	70	70
3	T3-T5	41	41	41	41	39	41
4	T4-T6	41	40	41	41	40	40
5	T5-T7	77	78	70	77	78	77
6	T6-T8	90	90	98	89	92	90
7	T7-T9	56	56	58	56	55	55
8	T8-T10	40	43	40	43	42	41
9	T9-T11	82	82	75	81	82	81
10	T10-T12	90	90	83	89	88	90
11	T11-T13	78	81	90	78	77	77
12	T12-T14	77	77	77	77	77	77

Implementation on Dataset-3

An arbitrary set of images of Saraswati College of Engg., Kharghar, Navi Mumbai were collected, using a panoramic set up. To capture the images the camera was mounted on a leveled tripod. While capturing the images camera was operated in manual mode where all camera parameters like aperture, shutter speed, focal length were constant. Since images are taken from different planes, the topology of mosaic was unknown. The motion between the images was unknown and was not assumed to be constant.



SCOE-1



SCOE-2



SCOE-3



SCOE-4



SCOE-5



SCOE-6



FIGURE 23: Sample images of SCOE and mosaic of images from SCOE1 to SCOE6

Observation

TABLE V : Location of maximum match

s. n.	Combination of images	Pixel based method		Fourier based method		Feature based method		Mutual Information based method		Principal component analysis	
		x-translation	y-translation	x-translation	y-translation	x-translation	y-translation	x-translation	y-translation	x-translation	y-translation
1	SCOE1 - SCOE2	157	-2	157	-5	156	-2	160	-2	159	-2
2	SCOE1 - SCOE2	163	35	164	35	163	35	163	35	164	35
3	SCOE1 - SCOE2	220	15	223	17	219	16	221	16	222	17
4	SCOE1 - SCOE2	163	10	161	5	162	10	161	09	162	10
5	SCOE1 - SCOE2	120	08	118	04	118	08	118	07	119	08

5. CONCLUSION

Various methods are reported in literature to register images which are in same band. In pixel based method cross correlation is used as similarity measure. It is observed that in natural images like buildings or scenery, correlation method shows match at multiple points. The feature based method makes use of features like point of intersection, edges, corners, centers of contours etc. for matching sample template with reference image. But this method is manual and hence time consuming. The method combining image features with correlation method have many advantageous properties of both feature-based and intensity based. It overcomes the limitation of intensity based method. Contour based methods do not use the gray values for matching and hence overcomes the limitations of spatial methods. Feature based method filter out the redundant information. Accuracy of this method is more but the limitation is, it is manual and slow. In frequency based method accuracy is more than correlation method but less as compared to other methods. But if we extract image features and then apply Fourier method accuracy increases. In frequency domain it should be noted that some form of interpolation must be used. These are some of the conclusions about methods used for registration of images which are in same spectral band. Image registration is difficult when images are obtained through different sensor types. Mutual Information, Hotelling Transform, Fuzzy logic are some of the approaches that can be used for multimodal image registration. For three different set of images SCOE1 to SCOE6, VR images form H1 to H15 and Infrared images track1 to track14 we carried out some of the popular algorithms and compared on the basis of point of match. Table I to Table V show the comparison. Combinational approach of wavelet and mutual information gives better

results as compared to wavelet - correlation combination and Fourier based image registration. Even wavelet mutual information combination can be used in case of multimodal image registration. If we involve mutual information as similarity criteria, speed decreases and hence we can not use this combination in applications where speed is required. If we use eigen values and eigenvectors for registration of images accuracy is more precise than that of other methods. We can use this method for template matching. It is observed that Hotelling Transform can be used to align region or boundaries with the eigenvectors of object but this method can be applied only to regions and boundaries. It is observed that method based on information theory yields a more accurate registration than any other registration method. But this method has its own limitations. When images are of low resolution, when images contain little information, or when the region of overlap is small then mutual information result in mis-registration. It has one more limitation of speed, when time is an important constraint one cannot use this method. Although it has some limitations entropy and mutual information are best approaches for multimodal image registration.

6. FUTURE SCOPE

A major challenge in the current literature is to perform population registration on large collections of data sets. Currently available tools pre selects a reference dataset that is template and registers in a pair wise fashion. The computational complexity and accuracy of this approach can be eliminated by performing a simultaneous registration on the whole population. The methods, we have explored in this paper have desirable computational speed for achieving population registration. An immediate next step would be to investigate this open problem that may lead to a significant contribution.

7. REFERENCES

- [1] Brown Gottesfeld L., "Survey of Image Registration Techniques", ACM Computing Surveys, 24, 4, 1992, 325-376.
- [2] Barbara Zitova, Jan Flusser, "Image registration methods:A survey", Image and Vision Computing 21(2003)977-1000.
- [3] J.B. Antoine Maintz and Max A. Vierger, "A Survey of Medical Image Registration", *Medical Image Analysis*,(1/98) volume 2. number1, pp 1-37
- [4] Leszek Chmielewski, Dorota Kozinska, "Image Registration", conf. Computer Pattern Recognition SystemsKOSYR 2003, May 26-29, 2001, Milkow,Poland, pp.163-168.
- [5] Mohammed ESSADIKI, "New technique for combining panchromatic and multispectral spot images for multipurpose image-maps".
- [6] Mert Rory Sabuncu, "Entropy-Based Image Registration". Phd Thesis.
- [7] S. K. Bose, "An Overview of Methods and Tools in Medical Image Registration.", 'National workshop on Digital signal processing' 3rd to 5th July, 2006.
- [8] J. Flusser, T. Suk, "Degraded Image Analysis: an Invariant Approach", IEEE transaction on Pattern Analysis and Machine Intelligence 20(1998)590-603.
- [9] J. Flusser, T. Suk, "A Moment Based Approach to Registration of Images with Affine Geometric Distortion", IEEE transactions on Geoscience and remote sensing 32(1994)382-387.
- [10] C. Zhang, "A Computational Vision Approach to Image Registration", IEEE Transaction on image processing, 2(3), pp. 311-326.

- [11] H. Manjunath, B. S. Mitra, 1995. "A contour-based Approach to Multisensor Image Registration", IEEE Transactions on image processing, 4(3), pp. 320-334.
- [12] Leszek Chmielewski, Dorota Kozinska , "Image Registration", conf. Computer Pattern Recognition SystemsKOSYR 2003, May 26-29, 2001, Milkow,Poland, pp.163-168.
- [13] Mert Rory Sabuncu,"Spatial Information in Entropy – Based Image Registration," in Biomedical Image Registration, LNCS 2717, Springer – Verlag, 2003.
- [14] Rafael C. Gonzalez, R. E. Woods, StevanL.Eddins, Digital Image Processing (Pearson Education, 2003).
- [15] Cahill, N. D. Williams, C.M. Shoupu, "Biomedical Imaging, Nano to Macro", 2006, 3rd IEEE International Symposium, issue, 6-9 April 2006, Page(s) 832-835.
- [16] J.P.W. Pluim, J.B.A. Maintz," Mutual Information Based Registration of Medical Images, Survey ", IEEE, Medical Imaging.
- [17] "Multimodal Image Registration using Generalized Survival Exponential Entropy" Medical image Computing, Springer, volume - 4191/2006, Cat Brain Analysis and registration, Pg-964-971, Friday, Sept 29, 2006.
- [18] Xiaoxiang Wang and Jie Tian, "Image Registration based on Maximization of Gradient Code Mutual Information ", Image Anal Streol 2005, 24:1-7.
- [19] Frederik Maes; Andre Collignon , "Multimodality Image Rregistration by Maximization of Mutual Information", IEEE Transactions on medical Imaging, vol.16, No. 2, April 1997.
- [20] J.P.Queiroz-Neto,M.F.M. Campos," Automatic Geometric and Radiometric Registration of Landsat, TM Images using Mutual Information".
- [21] C. Shannon,"A Mathematical Theory of Communication ", Bell System Technical Journal, vol 27, pp, 379-423, 623- 656, 1948.
- [22] Haim Schweitzer, "Optimal Eigenfeature Selection by Optimal Image Registration," cvpr,pp.1219, 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'99) - Volume 1, 1999.
- [23] Ma Debao Li Wugao Le Zhongxin Wang Jiefeng "The New Matrix Characteristic Methods of Image Fine Registration for Synthetic Aperture Radar Interferometry", Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International, 2000, Volume: 2, On page(s): 758-760, vol.2.
- [24] Haim Schweitzer," An Eigenspace Approach to Multiple Image Registration", NASA Goddard Space Flight Center, 1997, pages7—12.
- [25] Wen Cao; Bicheng Li; Yong Zhang," A Remote Sensing Image Fusion Method Based on PCA Transform and Wavelet Packet Transform", Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on Volume 2, Issue, 14-17 Dec. 2003 page(s): 976 - 981 Vol.2.
- [26] Yu-Te Wu¹, Takeo Kanade², Ching-Chung Li³ and Jeffrey Cohn⁴" Image Registration Using Wavelet-Based Motion Model", International Journal of Computer Vision, Springer Netherlands, 0920-5691 (Print) 1573-1405 (Online), volume 38, Number 2 / July, 2000, pages 129-152.

- [27] Tarek A. El-hazawi, Prachya , Chalermwat, Jacqueline Le Moigne, “*Wavelet-based Image Registration on Parallel Computers*” ,Proceedings of the 1997 ACM/IEEE conference on Supercomputing (CDROM), 1997, Pages: 1 - 9 .
- [28] Hala S. Own and Aboul Ella Hassanien, “*Q-shift Complex Wavelet-based Image Registration Algorithm*”, Springer Berlin / Heidelberg, 1615-3871 (Print) 1860-0794 (Online), Volume 30/2005, pages 403-410.
- [29] Azhar Quddus, Otman Basir, “*Wavelet-Based Medical Image Registration for Retrieval Applications*”, BMEI, vol. 2, pp.301-305, 2008 International Conference on BioMedical Engineering and Informatics, 2008.
- [30] Stone, Harold S.; Le Moigne, Jacqueline; McGuire, Morgan, “*Image Registration Using Wavelet Techniques*” , Proc. SPIE Vol. 3240, p. 116-125, 26th AIPR Workshop: Exploiting New Image Sources and Sensors, J. Michael Selander; Ed.
- [31] Ghazaw, A. Chalermwa, P.,” *Wavelet-Based Image Registration on Parallel Computers*”, Supercomputing, ACM/IEEE 1997 Conference, ISBN: 0-89791-985-8, On page(s): 20- 20.
- [32] Jiangsheng You, Weiguo Lu, Jian Li et. al. “*Image Matching for Translation Rotation and Uniform Scaling by Radon Transform*” 0-8186-8821-1/98, 1998 IEEE.
- [33] Tsao. J., “*Interpolation Artifacts in Multimodality Image Registration Based on Maximization of Mutual Informioatn* “, Medical imaging, IEEE volume 22, 7, July 2003 Page(s):854-864.

An Adaptive Two-level Filtering Technique for Noise Lines in Video Images

Baris Baykant ALAGOZ

*Department of Electrical and Electronics Engineering,
Inonu University, Malatya, Turkey*

baykant.alagoz@inonu.edu.tr

Mehmet Emin TAGLUK

*Department of Electrical and Electronics Engineering,
Inonu University, Malatya, Turkey*

mehmet.tagluk@inonu.edu.tr

Abstract

Due to narrow-band noise signals in transmission channels, visible lines of disturbance can appear in video images. In this paper, an adaptive method based on two-level filtering is proposed to enhance the visual quality of such images. In the first level, an adaptive orientation selective filter detects and clears the noisy lines in the image. In the second level, a median filter repairs defects resulting from the orientation selective filtering process and also filters the wide-band impulsive noise. It was observed that in case of periodic noisy lines in TV images, this filtering technique can sufficiently enhance the image quality and improve the SNR level.

Keywords: Adaptive noise filter, Wireless video image enhancement.

1. INTRODUCTION

In long distance wireless video transmission systems, periodic noise line patterns commonly appear on the received image. Such noise line patterns in wireless transmission are mostly caused from long-term narrow-band signal interference to the communication channel. In many cases, such interference of by noise signals in the channel is unavoidable and therefore the removal of these noise signals has to be carried out on the received image by using filtering techniques. A basic notch reject filter has been applied for the removal of the noise lines on the images received from the long-distance space missions [1]. Nowadays, with increasing usage of wireless video transmission systems in day-to-day applications there is an increasing demand to develop such filtering techniques to work on received images.

In practice, unlicensed transmitter interference, noise in electronics, multi-path effects, loss of horizontal or vertical synchronization are all seen to cause periodic noise lines on the received image. These noise lines can severely mislead computer vision algorithms employed in autonomous remote control systems used in unmanned vehicles [2-5].

Periodic noise lines in images are commonly seen in imaging systems, which use a row scanning mechanism in the construction of the image data when a long-term noise signal affects the system. For example, mechanical and acoustic vibrations in force sensors were seen to decrease the signal to noise ratio (SNR) of images scanned by an Atomic Force Microscope (AFM) at video rate [6].

In the image enhancement field, a variety of methods have been developed to filter out the effects of random noise on images. Most of these studies have been focused on preserving singular features of the image such as edges, while smoothing other segments of the image [7-11]. Alternatively a nonlinear adaptive filter based on a neural network has been proposed for reducing the additive noise [12]. Many adaptive image restoration methods that analyse the noise and optimise the behaviour of the filter to improve overall filtering performance [13-15]. Specifically, for digital TVs the filters compromise edge detection, and an automated modification of filter coefficients has been addressed in detail by Chan et al. [7, 10]. These filters were mainly

developed to deal with random additive noise in images. In their mathematical formulation, the received signal has been defined as:

$$u_r(x) = u(x) + n(x), \quad (1)$$

where, $n(x)$ is the random noise and $u(x)$ is the original noise-free image. Due to the fact that the noise components are treated as a part of the image signal $u(x)$, such filters designed for removing random noise are not able to effectively deal with narrow-band noise lines.

In our study, we assume that the received signal has a narrow-band noise signal besides a random noise signal. Under this assumption, a noise signal can be modelled as:

$$u_r(x) = u(x) + n_n(x) + n(x) \quad (2)$$

where, $n_n(x)$ represents the narrow-band noise.

To enhance the received image coming from highly noisy channel modelled by equation (2), we proposed a filtering structure which is composed of a frequency domain orientation selective notch reject filter for the elimination of the narrow-band noise ($n_n(x)$) [1, 16] and a spatial median filter for impulsive noise [8]. The block diagram of such a two-level filtering is presented in Figure 1.

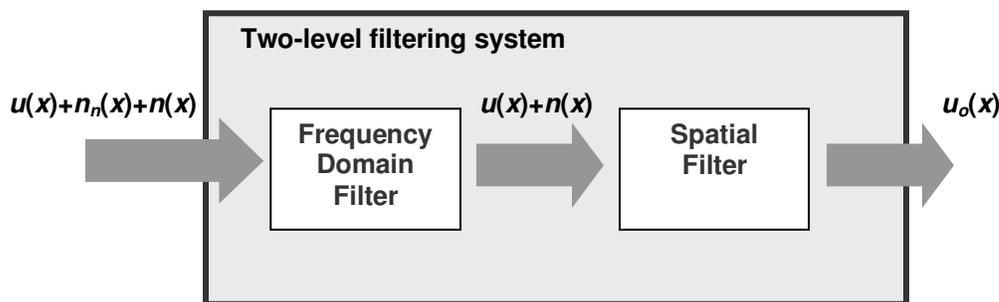


FIGURE 1: Block diagram of two-level filtering.

A noisy image is first transformed to red, green and blue color channels, each color channels being solely filtered by the two-level filter designed. In order to adapt the angle of notch reject band to the slope of the noise lines in the received images, the proposed orientation selective filter first detects the spectral region belonging to noise lines, via measuring the local spectral power density, and then it orients the reject band onto the region where the power of noise lines is intensified.

2. PROPOSED METHOD

Frequency domain filtering [16] usually provides a good performance in the filtering of narrow-band noise signals, because of allow us eliminating frequency components in a narrow frequency band. Basic design methodology of a frequency domain filter is as follows: The spatial two-dimensional image data are first transformed into frequency components via a two-dimensional Fast Fourier Transform (2DFFT) and then a two-dimensional mask is applied to suppress undesired frequency components, and finally an inverse Fast Fourier Transform (2DIFFT) is used to obtain the filtered image in the spatial domain, as illustrated in Figure 2.

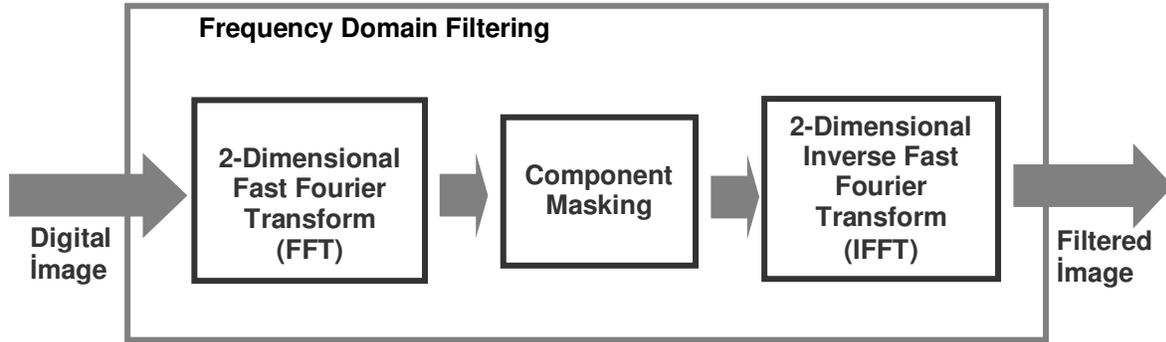


FIGURE 2. Frequency domain filtering with component masking

Masking is done by multiplication of each mask element with elements of R , G and B image as follows:

$$\begin{aligned}
 R_{fm}(u, v) &= R_f(u, v) \cdot Mk(u, v), \\
 G_{fm}(u, v) &= G_f(u, v) \cdot Mk(u, v), \\
 B_{fm}(u, v) &= B_f(u, v) \cdot Mk(u, v),
 \end{aligned}
 \tag{3}$$

where, R_f , G_f and B_f are the FFT of the R , G and B matrix, R_{fm} and G_{fm} and B_{fm} are the masked frequency components of the RGB image. The mask matrix Mk has real value, $Mk(u, v) \in R$, in such cases this mask takes effect on the amplitude of the spectral components of images. The phases of frequency components are preserved in this masking operation.

In this study, we introduce a mask Mk_ϕ generation function that provides a directional stop-band over high frequencies to suppress periodic line patterns in the image for a given ϕ angle. The proposed mask is formed from superposition a suppressing channel passing through the origin of the 2D spectrum and a directional pass band at the origin.

The construction of the mask is as follows. The suppressing channel, whose direction is controlled by a given angle, ϕ , was constructed on a base line passing through the origin

$(\frac{M}{2}, \frac{N}{2})$ as presented in Figure 3.

$$u = -\tan(\phi) \cdot (v - \frac{N}{2}) + \frac{M}{2},
 \tag{4}$$

where M and N determine the size of the mask, Mk_ϕ , and ϕ is angle of the base line. The shortest distance of any point (u, v) to this base line, denoted by $d_s(u, v, \phi)$, was derived as:

$$d_s(u, v, \phi) = \begin{cases} \sqrt{(u - u_b(u, v, \phi))^2 + (v - v_b(u, v, \phi))^2} & , 90^0 < \phi < 270^0 \\ \left| v - \frac{N}{2} \right| & , \phi = 90^0 \\ \left| u - \frac{M}{2} \right| & , \phi = 270^0 \end{cases} \quad (5)$$

where $(u_b(u, v, \phi), v_b(u, v, \phi))$ is the nearest point of the base line to the point (u, v) on the mask, and calculated by the following expressions:

$$u_b(u, v, \phi) = \tan(\phi) \cdot \left(v_b(u, v, \phi) - \frac{N}{2} \right) + \frac{M}{2}. \quad (6)$$

$$v_b(u, v, \phi) = \frac{\tan^2(\phi) \cdot N + 2 \cdot v - \tan(\phi) \cdot M + 2 \cdot \tan(\phi) \cdot u}{2 \cdot (\tan^2(\phi) + 1)}. \quad (7)$$

The shortest distance of any point (u, v) to the centre of the mask, denoted by $d_c(u, v)$, is calculated as:

$$d_c(u, v) = \sqrt{\left(u - \frac{M}{2}\right)^2 + \left(v - \frac{N}{2}\right)^2}. \quad (8)$$

The mask generation function is then formed as:

$$Mk_\phi(u, v) = \left(1 - \exp\left(-\frac{d_s(u, v, \phi)}{\alpha_s}\right)\right) + \exp\left(-\frac{d_c(u, v, \phi)}{\alpha_c}\right) \cdot \exp\left(-\frac{d_s(u, v, \phi)}{\alpha_s}\right), \quad (9)$$

or in a more compact form:

$$Mk_\phi(u, v) = 1 + \left(\exp\left(-\frac{d_c(u, v, \phi)}{\alpha_c}\right) - 1\right) \cdot \exp\left(-\frac{d_s(u, v, \phi)}{\alpha_s}\right), \quad (10)$$

where α_s and α_c are the standard deviations determining the width of directional suppression channel. These parameters also control smoothness of the transition from the stop to pass band of the filter. This smooth transition consequently eliminates ringing effects appearing around patterns in the image. In Figure 4, Mk matrix generated for various ϕ are illustrated.

The angle ϕ is adaptively determined in a range of $[\phi_{\min}, \phi_{\max}]$ by applying the following steps:

Step 1: Calculate $Pw(\phi)$ for all ϕ angles from ϕ_{\min} to ϕ_{\max} . ($Pw(\phi)$, expressed by equation (11), is the average spectral power in the image spectrum that is effected by the suppression channel of Mk_ϕ)

Step 2: Use the Mk_ϕ mask in filtering for the ϕ , at which $Pw(\phi)$ is a maximum and value of this maximum exceeds a predefined filter activation threshold (P_{th}). In the case, all $Pw(\phi)$ $\phi \in [\phi_{\min}, \phi_{\max}]$ are lower than P_{th} , the current image can bypass filtering. This implies that the noise lines do not have enough power to activate filtering.

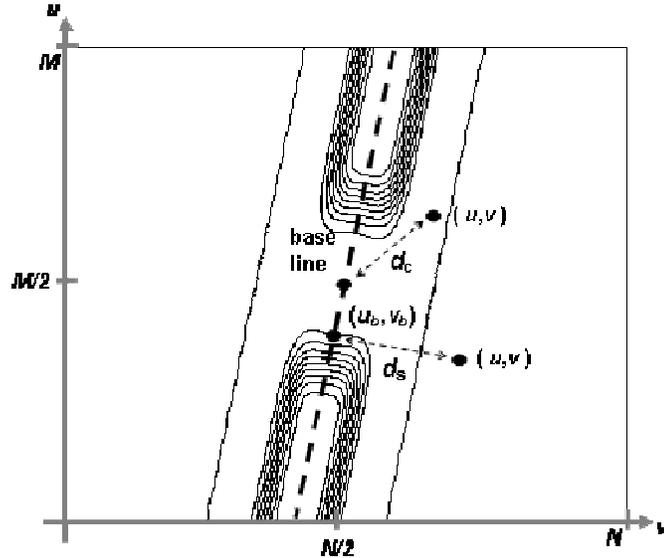


FIGURE 3. A contour plot of a suppression channel and its relevant parameters with respect to a base line passing through the centre of the spectrum.

The average image power in the suppression channel $P_w(\phi)$ can be calculated by:

$$P_w(\phi) = \frac{1}{\sum_{u=0}^{M-1} \sum_{v=0}^{N-1} (1 - Mk(u, v, \phi))} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} (1 - Mk(u, v, \phi)) \cdot (|R_f(u, v)|^2 + |G_f(u, v)|^2 + |B_f(u, v)|^2) \quad (11)$$

The term $(1 - Mk(u, v, \phi))$ in equation (11) is the weighting function for spectral components. Equation (11) is a discrete two dimensional extension of a single variable weighted average spectral power given as:

$$P_w = \frac{1}{\int w(f) df} \int w(f) \cdot |I(f)|^2 df \quad \text{for a weight function } w(f).$$

The choice of α_c and α_s parameters are in fact based on trial and error. However it was observed that, in this particular application, for TV images with a size of 640x480 pixel, then $\alpha_c = 2828$ and $\alpha_s = 316$ can give a satisfactory results.

After removing the noise line patterns from the image by adaptive frequency domain filtering, a 3x3 median filter is applied to the image in order to improve the image quality. This median filtering reduces the deformations resulting from the removal of noise line patterns, as well as the impulsive noise on the filtered image.

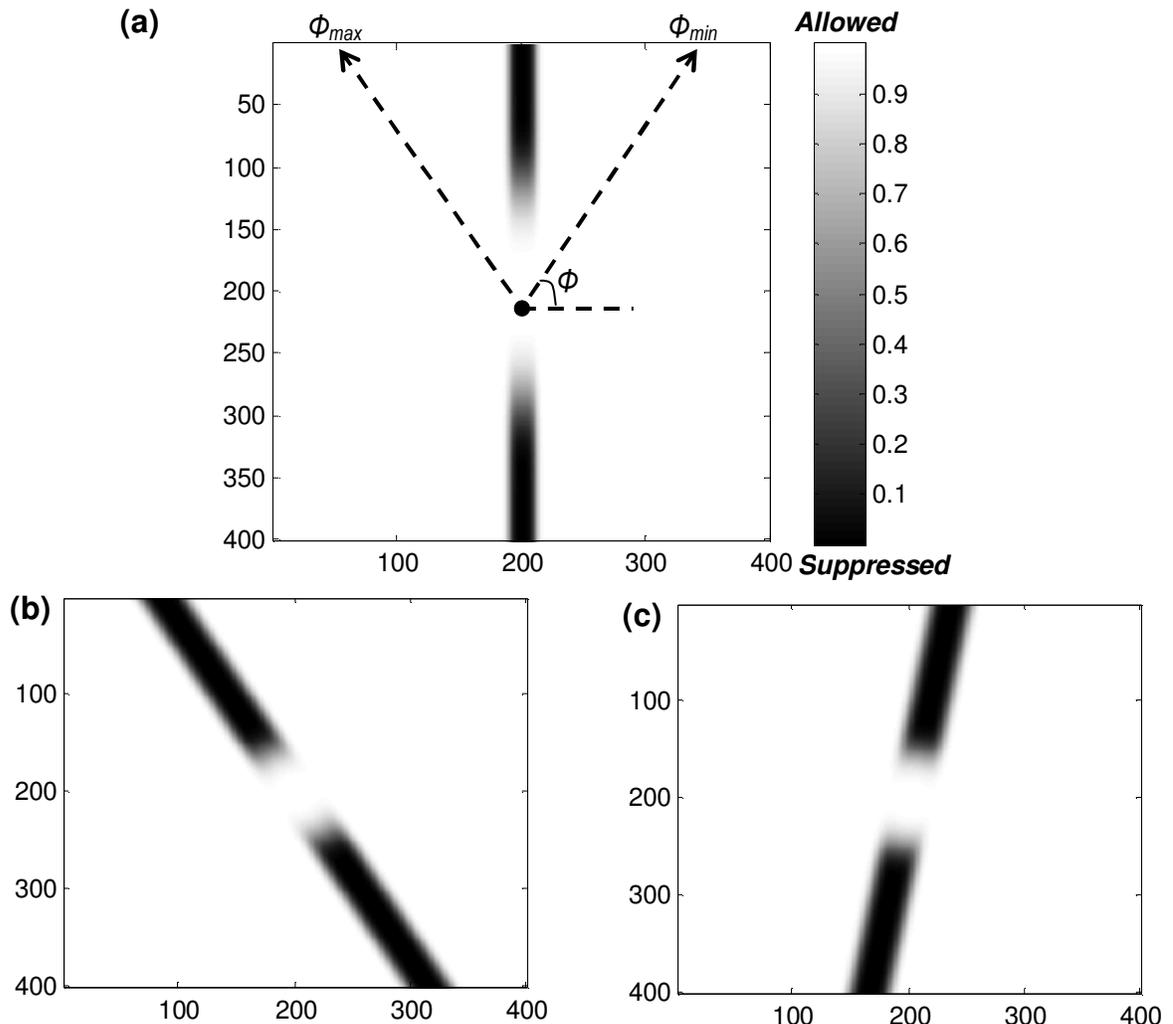


FIGURE 4. (a) Mk matrix generated for $\phi = 90^0$ (b) Mk matrix generated for $\phi = 120^0$ and (c) Mk matrix generated for $\phi = 80^0$

3. EXPERIMENTAL RESULTS

In Figure 5 and in Figure 6, TV images from a real media broadcast captured by a commercial TV Card were enhanced by the adaptive two-level filtering process described in the previous sections. The calculated spectral power as a function angle ϕ and the resulting Mk mask are given in the sub-figures (c) and (d), respectively.

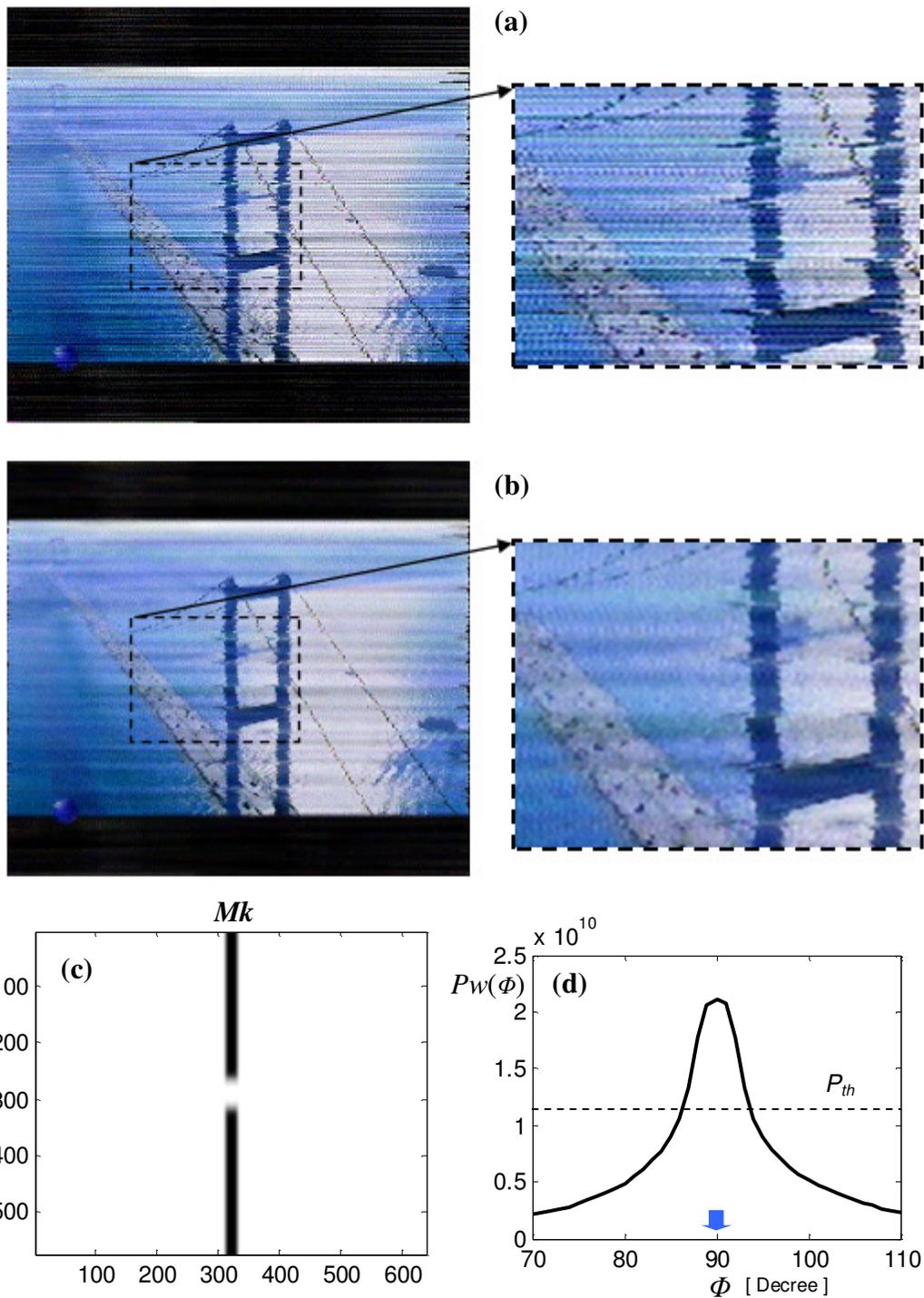


FIGURE 5. (a) Noise test image from TV, (b) Enhanced image by filter system ($\alpha_c = 2828$, $\alpha_s = 316$), (c) Generated Mk matrix, and (d) $P_w(\phi)$ values in the adaptation process. The filter applied a mask for $\phi = 90^0$.

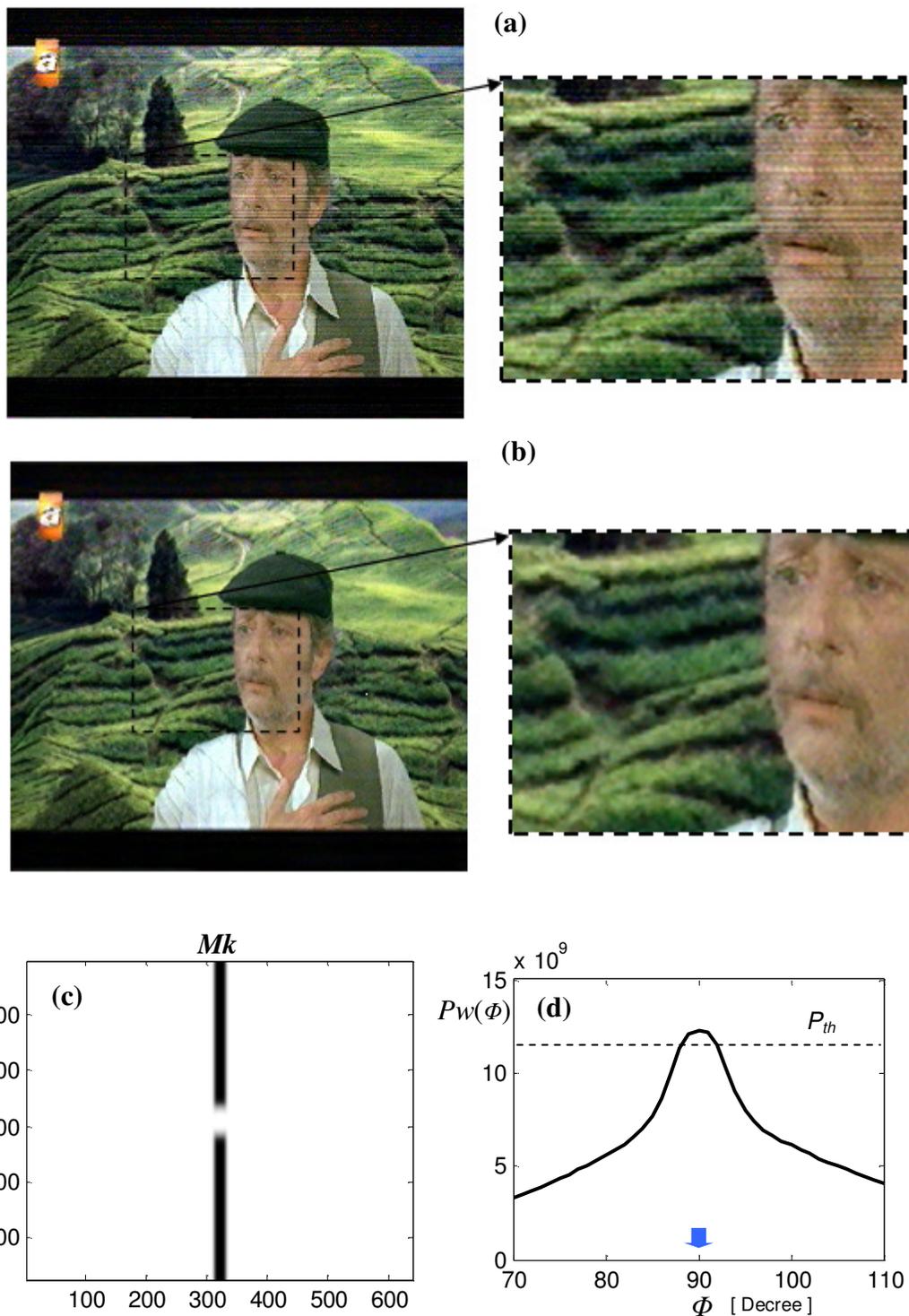


FIGURE 6. (a) Noise test image from TV, (b) Enhanced image by filter system ($\alpha_c = 2828$, $\alpha_s = 316$) , (c) Generated Mk matrix, and (d) $Pw(\phi)$ values in the adaptation process. The filter applied a mask for $\phi = 90^0$.

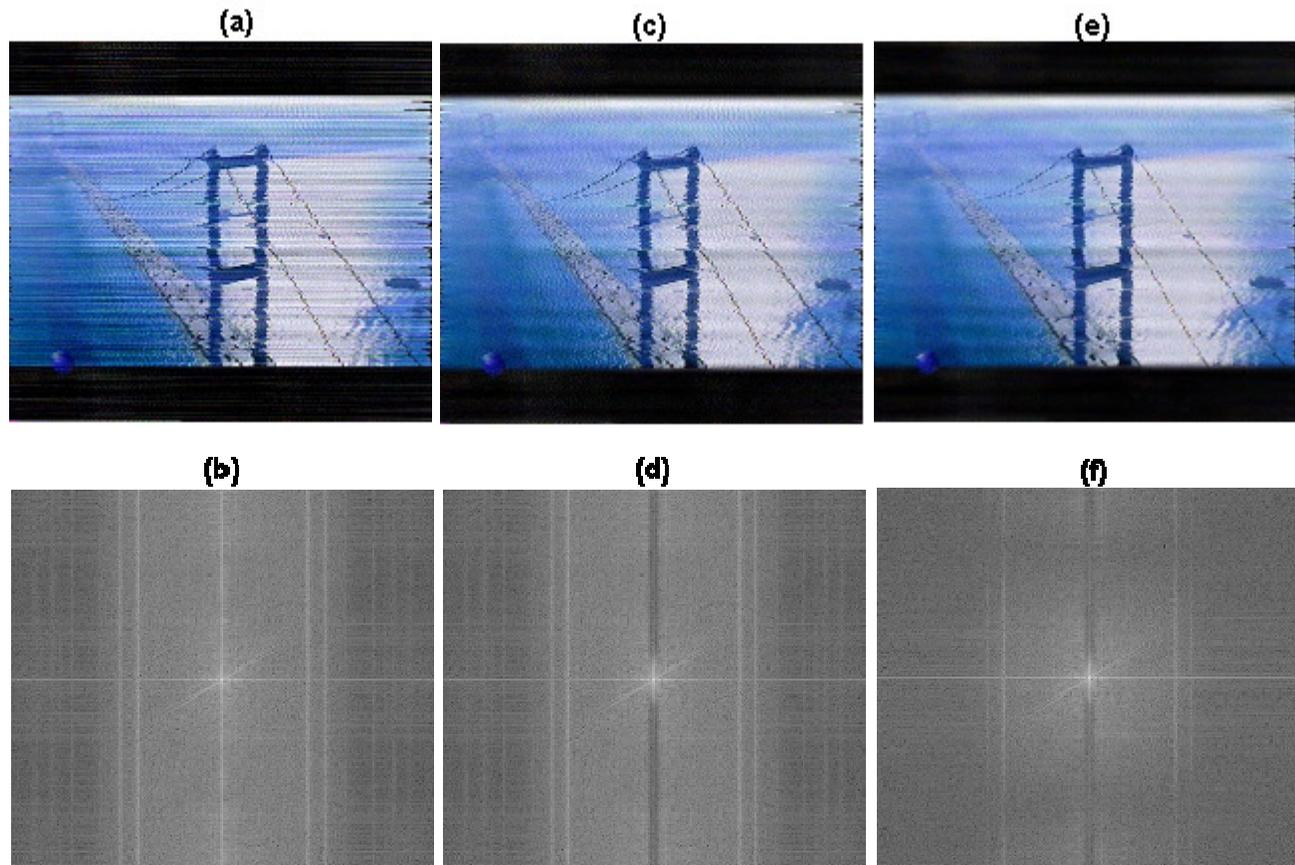


FIGURE 7. (a) and (b) are noise test images and their frequency spectrums , (c) and (d) are adaptive orientation selective filtering results and their frequency spectrums , (e) and (f) are median filtered images of (c) and their frequency spectrums..

In Figure 7, the effects of filtering at each level on the test image and its frequency spectrum are illustrated. After cleaning noise line patterns from the image, as seen in Figure 7(c), median filtering with 3x3 windowing can be seen to effectively restore the image from impulsive noise and deformation resulting from the frequency filtering. Median value with respect to image data in 3x3 windows have an effect of repairing notch filtering that is a vertical suppression band passing through the centre in Figure 7(d). This effect can be seen as strengthening some of the frequency components in the vertical suppression band in Figure 7(f).

4. TEST ON SYNTHETIC IMAGE AND SNR LEVELS

In this section, the noisy television test image shown in Figure 8(a) was distorted by a synthetic noise line pattern. This additive noise was generated by a sinusoidal signal with an amplitude of 80 and a frequency set to 0.001, as in Figure 8(c), and a pseudo-random noise signal with a uniform distribution in the interval $(-50,50)$, as in Figure 8(d). After adding these noise signals to the television test image in accordance with Equation 2, a noise test image was obtained as shown in Figure 8(b).

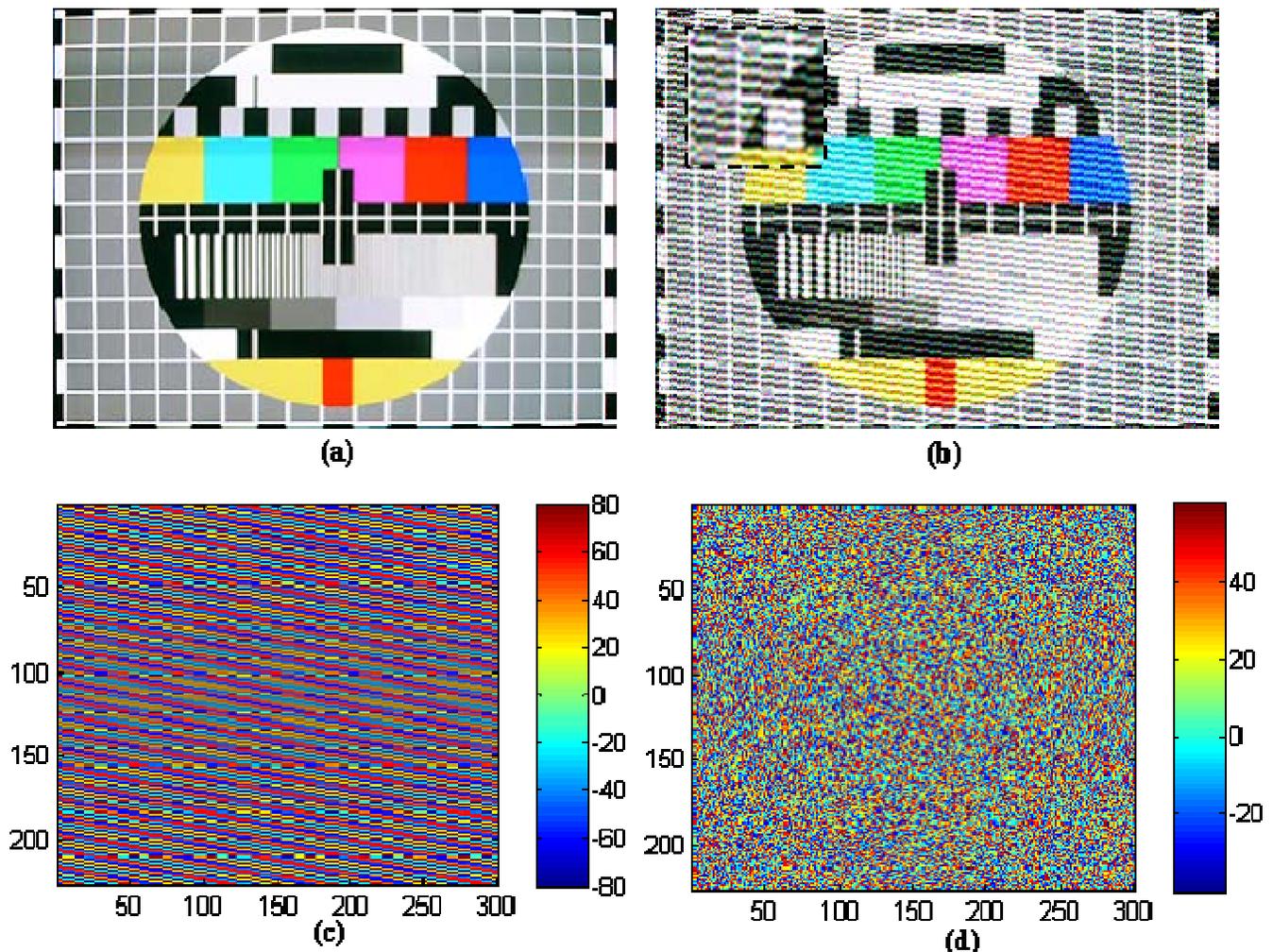


FIGURE 8. (a) and (b) are noise test images and their frequency spectrums , (c) and (d) are adaptive orientation selective filtering results and their frequency spectrums , (e) and (f) are median filtered images of (c) and their frequency spectrums.

The noisy television test image given in Figure 8(b) was enhanced by using three different basic filters and the proposed two-level filtering, and the results obtained are demonstrated in Figure 9. In Figure 9 it can be seen that the median-filter and Wiener filter with window size of 3x3 did not sufficiently remove the line patterns in the test image. Even though the Gaussian low pass filter blurred the image, the noise lines are still visible. The proposed method detected the noise lines at $\phi = 83^0$ and adapted the orientation selective filter according to the noise lines. After the application of both orientation selective filter in the frequency domain and the median filter in the spatial domain the noise lines were removed without too much blurring of the image as seen in Figure 9(d).

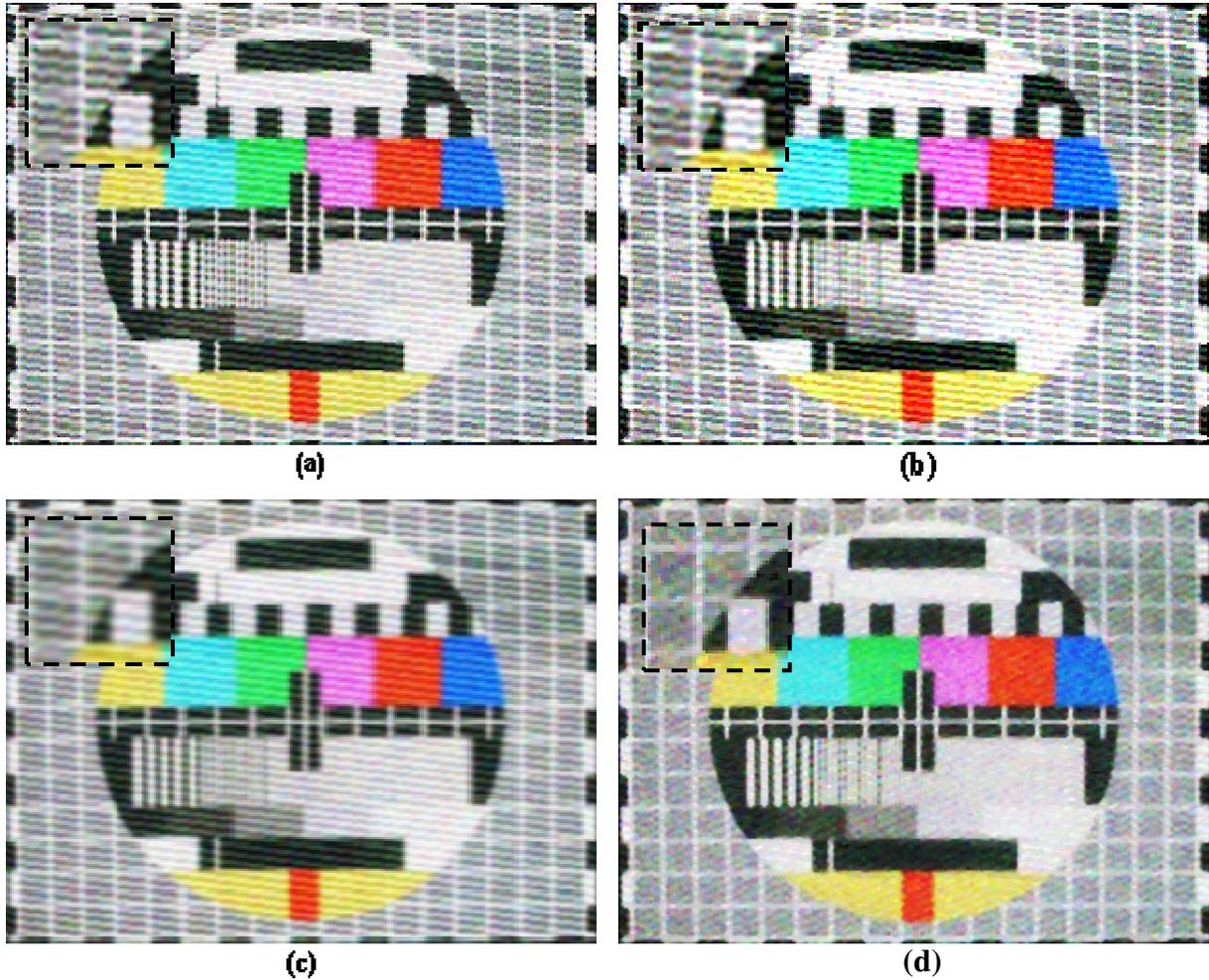


Figure 9. (a).Median filtered image, (b) Wiener filtered image, (c) Low pass filtered image by a Gaussian function mask, (d) The proposed two-level filtered image.

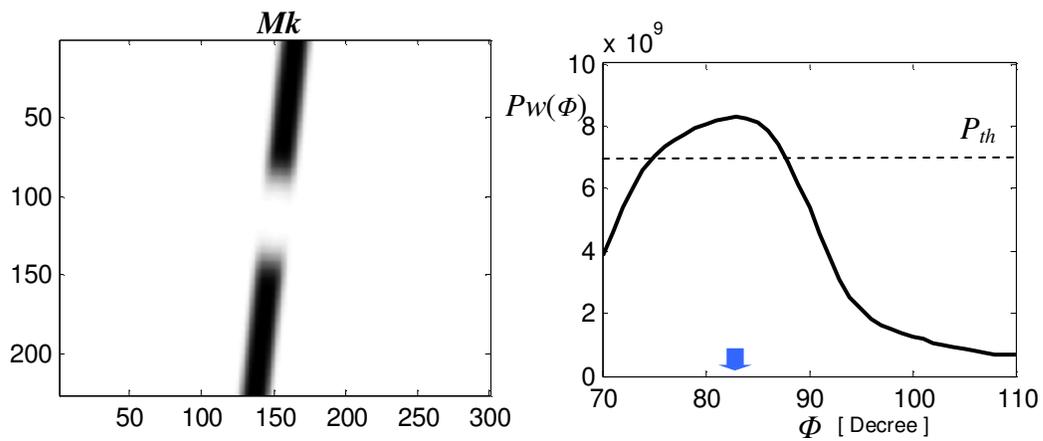


FIGURE 10. Generated Mk matrix and $Pw(\phi)$ values in the adaptation process. The filter was found to be best when $\phi = 83^0$.

In Figure 11, we present Signal to Noise Ratio (SNR) of these filters on the noisy television test image. In these tests, the P_{th} of the orientation selective filter activation was set to 3.10^9 . In this case, when the SNR of the received image is over 40 dB, the noise level remains below P_{th} and the orientation selective filtering was bypassed and the adaptive two level filtering performed the median filtering on the received image, solely. When the SNR of the test image was decreased to lower levels, the proposed two-level filtering detected the visible noise line patterns in the received image and removed them from the image. In this way, it maintains higher SNRs in the case of a video transmission in highly noisy channel.

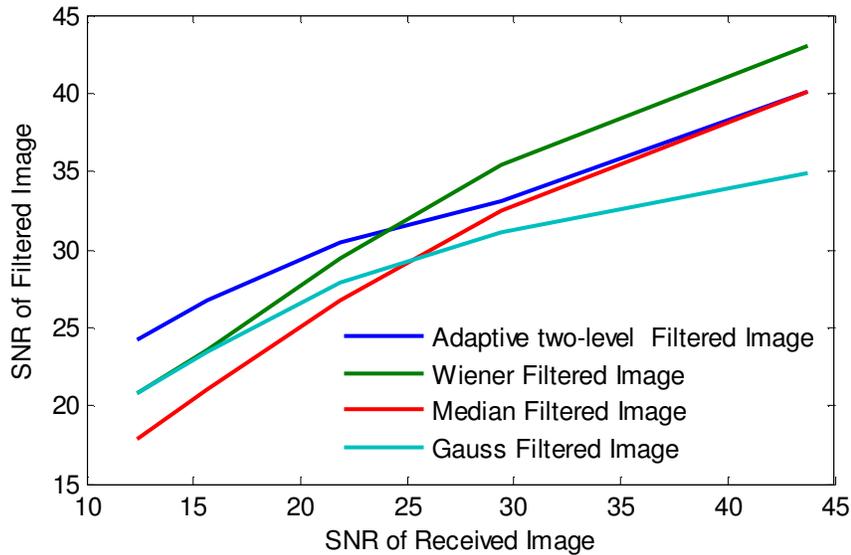


FIGURE 11. SNR of the filtered images for various additive noise levels.

5. CONCLUSIONS

The noise channel model given by equation (2) provides a more realistic model of today's communication channels. In this manner, a more adaptive and sophisticated filtering system is needed to restore video images. The proposed adaptive two-level filtering method, which integrates an adaptive orientation selective filter and a median filter, was seen to effectively enhance video images containing visible periodic noise lines.

The results obtained show us that the narrow-band noise signal can be detected and filtered by mask adaptive frequency domain filtering and the slight deformations resulting from such frequency domain filtering on the image data can be repaired by the median filtering.

6. REFERENCES

- [1] R.C. Gonzalez, R.E. Woods, S.L. Eddins. Digital image processing using MATLAB. New Jersey:Prentice Hall, 2004.
- [2] J. Byrne, R. Mehra." Wireless video noise classification for micro air vehicles." Proceedings of the 2008 Association for Unmanned Vehicle Systems International (AUVSI) Conference, 2008, pp.1-25.

- [3] P.Y. Oh, W.E. Green, G. Barrows. "Flying insect inspired vision for autonomous aerial robot maneuvers in near-earth environments." IEEE International Conference on Robotics and Automation (ICRA) , 2004, pp.2347–2352.
- [4] J. Byrne, M. Cosgrove, R. Mehra. "Stereo based obstacle detection for an unmanned air vehicle." IEEE International Conference on Robotics and Automation (ICRA'06), 2006.
- [5] J. Byrne, C.J. Taylor. "Expansion segmentation for visual collision detection and Estimation." IEEE International Conference on Robotics and Automation Kobe International Conference Center, 2009, pp.875-882.
- [6] G. Schitter, M.J. Rost. "Scanning probe microscopy at video-rate." Materials Today, vol.11, pp.40–48, 2008.
- [7] T.F. Chan, S. Osher, J. Shen. "The digital TV filter and nonlinear denoising." Image Processing IEEE Transaction, vol.10, pp.231-241, 2001.
- [8] C.A. Glasbey, G.W. Horgan. Image analysis for the biological science. Statistics in practice, New York:John Wiley & Sons, 1995.
- [9] B. Jaehne. Digital image processing, Concepts, algorithms, and scientific applications, Berlin:Springer, 1997.
- [10] S. Osher, J. Shen. Digitized PDE method for data restoration, In: G. A. Anastassiou (Ed.), Analytic-computational methods in applied mathematics, Chapman & Hall/CRC, 2000.
- [11] A. Lev, S.W. Zucker, A. Rosenfeld. "Iterative enhancement of noisy images." IEEE Transactions on Systems, Man and Cybernetics, vol.7, pp.435-443, 1977.
- [12] L. Corbalan, G.O. Massa, C. Russo, L. Lanzarini, A. De Giusti. "Image recovery using a new nonlinear adaptive filter based on neural networks." Journal of Computing and Information Technology CIT, vol.14, pp.315–320, 2006.
- [13] H.S. Wong, L. Guan. "A neural learning approach for adaptive image restoration using a fuzzy model-based network architecture." Neural Networks IEEE Transactions, vol.12, pp.516–531, 2001.
- [14] E.S. Hore, B. Qiu, H.R. Wu. "Adaptive noise detection for image restoration with a multiple window configuration." Image Processing Proceedings. 2002 International Conference, 2002, pp. 329-332.
- [15] F. Russo. "A method for estimation and filtering of Gaussian noise in images." Instrumentation and Measurement, IEEE Transactions, vol.52, pp.1148–1154, 2003.
- [16] R.P. Matei. "Design method for orientation-selective CNN filters." Circuits and Systems, Proceedings of the 2004 International Symposium vol.3, 2004, pp.105-8.

Fast Complex Gabor Wavelet Based Palmprint Authentication

Jyoti Malik

*Electrical Department
National Institute of Technology
Kurukshetra, 136119, India*

jyoti_reck@yahoo.com

Ratna Dahiya

*Electrical Department
National Institute of Technology
Kurukshetra, 136119, India*

ratna_dahiya@yahoo.co.in

G Sainarayanan

*Maples ESM Technologies
Chennai, India*

sai.jgk@gmail.com

Abstract

A biometric system is a pattern recognition system that recognizes a person on the basis of the physiological or behavioral characteristics that the person possesses. There is increasing interest of researchers in the development of fast and accurate personal recognition systems. In this paper, Sliding window method is used to make the system fast by reducing the matching time. The reduction in computation time indirectly reduces the overall comparison time that makes the system fast. Here, 2-D Complex Gabor Wavelet method is used to extract features from palmprint. The extracted features are stored in a feature vector and matched by hamming distance similarity measurement using sliding window approach. Reduction of 74.12% and 90.32% in comparison time is achieved using Sliding window methods. The improvement in time is indicated by experimental results that makes a system rapid.

Keywords: Palmprint Authentication, Complex Gabor Wavelet, Similarity Measurement, Sliding Window Method.

1. INTRODUCTION

Biometric identification of a person by his/her physiological or behavioral characteristics, like face, finger, palmprint, gait, signature, voice etc. has become increasingly popular in modern personal identification and verification systems [1-3]. In this paper, palmprint biometric is selected for personal authentication as it is unique and relatively low resolution images (less than 100 dpi) are sufficient to extract its unique features [4][5].

Palmprint features include line features, geometry features, point features, texture features and statistical features. In this paper, line features are extracted using Complex Gabor Wavelet Transform method. A complex Gabor wavelet is defined as the product of a Gaussian kernel times a complex sinusoid. The line features extracted by complex gabor wavelet [6-15] at various values of theta is stored in the feature vector. The feature vector is matched by Hamming Distance similarity measurement using sliding window method.

In this paper, the palmprint authentication system is divided in following two subsystems:

- a) Pre- Authentication System
- b) Authentication System

In Pre-authentication system, a database of Gabor-Palmprint features is prepared. Reference threshold values are also identified and stored in database. These values will be later used by Authentication system.

In Authentication system, the authenticity of a person being genuine or imposter is identified with the help of Reference threshold values stored in Pre-authentication system database.

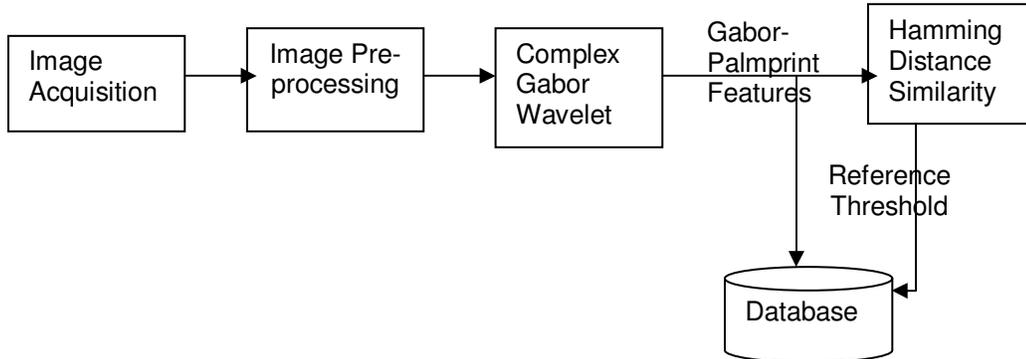


FIGURE 1: Palmprint Pre-Authentication system

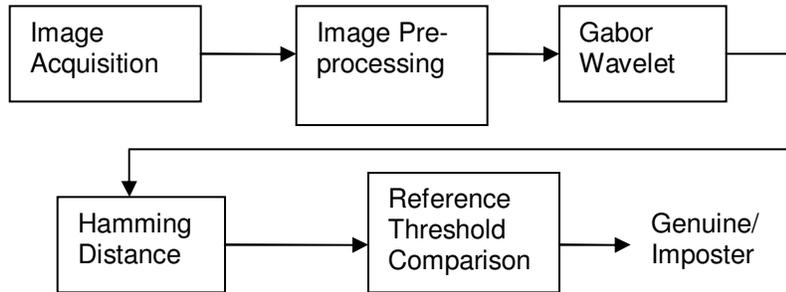


FIGURE 2: Palmprint Authentication System

In Palmprint-based personal recognition system matching time plays very important role in making the real time authentication system. In this paper, we have proposed fast palmprint authentication system. The reduction in matching time is done using Sliding window methods and explained later in section 4.

In the following sections, Section 2 defines palmprint feature extraction by our proposed Complex gabor wavelet method. Section 3 explains the feature matching method by hamming distance using sliding window method. Section 4 explains Sliding window method (SWM), SWM1, SWM2 and Section 5 defines the reference threshold calculation. Section 6 discusses the experimental setup with the results. Section 7 concludes the conclusion.

2. FEATURE EXTRACTION

The desired line features are extracted from the palmprint using Complex Gabor wavelet method which extracts line features from the input palm-print image. The Gabor wavelet is basically a Gaussian (with variances s_x and s_y along x and y -axes respectively) modulated by a complex sinusoid (with centre frequencies U and V along x and y -axes respectively) described by the following equation:

$$G_i(x, y) = \frac{1}{2 * \pi i * s_x * s_y} * \exp \left(\left[\frac{-1}{2} \left\{ \left(\frac{x}{s_x} \right)^2 + \left(\frac{y}{s_y} \right)^2 \right\} \right] \right) * M_i(x, y, f) \quad (1)$$

$$i = 1, 2$$

$$M_1(x, y, f) = \cos\left[2 * \pi i * f * \sqrt{(x^2 + y^2)}\right]$$

$$M_2(x, y, f) = \cos\left[2 * \pi i * f * (x * \cos \theta + y * \sin \theta)\right]$$

Where s_x and s_y are the variances along x and y -axes respectively, f is the frequency of the sinusoidal function, θ is the orientation of Gabor wavelet, G_1 and G_2 are the output Gabor wavelets. The sample of Gabor wavelet convolution with the palmprint image is shown in Fig. 3.

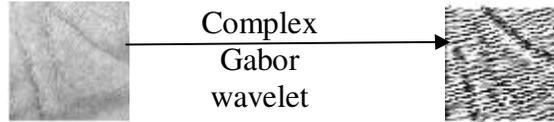


FIGURE 3: Feature extraction by Complex Gabor Wavelet

The feature vector contains features extracted by Complex Gabor wavelet at different orientations i.e. $0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$ and 180° as shown by figure 4.

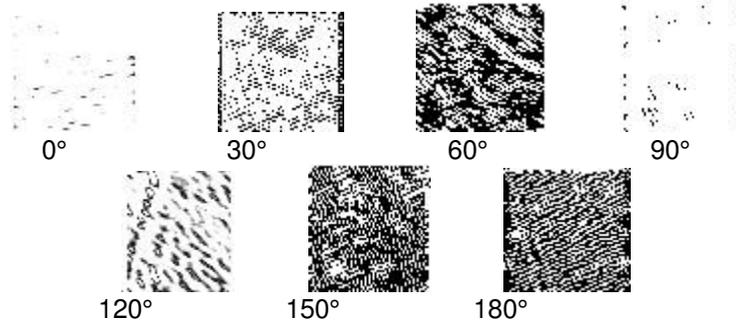


FIGURE 4: Complex Gabor wavelet features images

The feature vector contains 7 elements at corresponding to each orientation. The Feature vector matrix is given by (2)

$$CGWF = [CGWF_0, CGWF_1, CGWF_2, CGWF_3, CGWF_4, CGWF_5, CGWF_6] \quad (2)$$

3. FEATURE MATCHING BY HAMMING DISTANCE AND SLIDING WINDOW METHOD

The complex gabor wavelet line feature vectors (line information) are matched by Hamming distance similarity measurement method but firstly the line information (complex gabor wavelet features) extracted is binarized by the following equation (3):

$$CGWF_n(i, j) = \begin{cases} 1, & CGWF_n(i, j) > 0 \\ 0, & CGWF_n(i, j) \leq 0 \end{cases} \quad (3)$$

where, $CGWF_n(i, j)$ = complex gabor wavelet features corresponding to n^{th} orientation, $n = 0, 1, 2, \dots, 6$, i and j are the rows and columns of the complex gabor wavelet features.

Hamming Distance calculates the difference between two binary feature vectors using EX-OR operation and can be defined as in (4):

$$HD_n = \sum_i \sum_j (FV(i, j) \oplus FV_{DB}(i, j)) \quad (4)$$

where, HD_n denotes the Hamming distance at orientation n , $n = 0, 1, 2, \dots, 6$, i and j is the row and column of the complex gabor wavelet feature vector, \oplus is the exclusive OR operation, FV denotes the feature vector of the person to be matched, FV_{DB} denotes the feature vector in database. Here, FV is same as $CGWF$.

The hamming distance value is calculated by using sliding window method. In sliding window method, the ROI is reduced by the window size (WS) and the window of $((60-WS) \times (60-WS))$ slides over the rows and columns out of 60×60 pixels considered for Hamming distance matching. The minimum value of the hamming distance values is considered. Fig. 5 shows the sliding window approach.

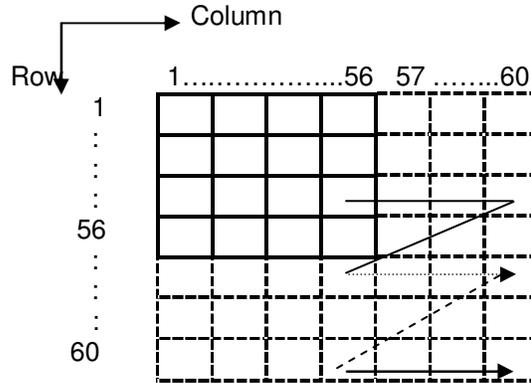


FIGURE 5: Sliding Window Approach with window size 4 and palmprint size 60×60

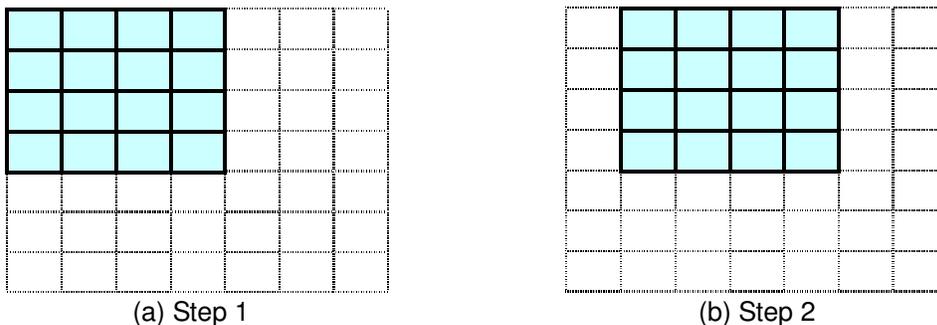
The modified Hamming distance value at n orientation with window size WS is defined in (5) as:

$$HD_{WSn} = \sum_i^{60-WS} \sum_j^{60-WS} (FV(i, j) \oplus FV_{DB}(i, j)) \quad (5)$$

where, HD_{WSn} denotes the Hamming distance with window size WS and at an orientation n , $n = 0, 1, 2, \dots, 6$, i and j is the row and column of the Complex gabor wavelet feature vector, \oplus is the exclusive OR operation, FV denotes the feature vector of the person to be matched, FV_{DB} denotes the feature vector in database. The minimum value out of 16 (For window size $WS=4$, $4 \times 4 = 16$) values of Hamming distances is chosen as Hamming distance as calculated in (6).

$$HD_n = \min(HD_{1-n}, HD_{2-n}, HD_{3-n}, \dots, HD_{16-n}) \quad (6)$$

The various steps in sliding window method are shown in figure 6.



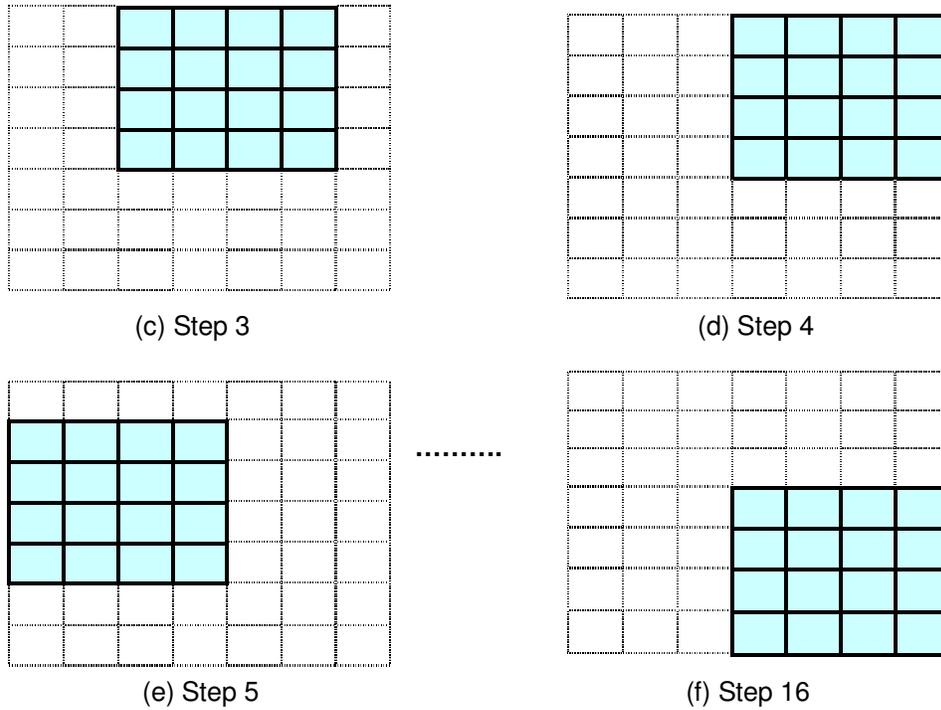


FIGURE 6: Various steps in sliding window method

The average of all Hamming distances for n orientations is calculated as shown in (7).

$$AHD = \left(\frac{\sum_{n=0,1,2,\dots,6} HD_n}{n} \right) \tag{7}$$

where, HD_n denotes the hamming distance for n^{th} orientation, AHD denotes the average value of hamming distance. The average value will help in finding the reference threshold value.

If, the Hamming distance value of two feature vectors is less than reference threshold value, feature vectors are considered to be from same hands otherwise different hands.

4. COMPARISON TIME IMPROVEMENT USING FAST SLIDING WINDOW METHOD

The sliding window method is the basic approach of feature matching and an accurate method but slow because of window size. The window size is chosen in such a manner to avoid alignment problem in the palmprint images. Here, two sliding window approaches are proposed. Sliding window method 1 and Sliding window method 2 is about using palmprint segment in such a manner so that the matching operation time is reduced. In this paper, complexity of the algorithm of SWM is reduced in SWM1 and SWM2.

4.1 Sliding Window Method

SWM is the basic technique used for matching two feature vectors. In feature matching, full image is matched with the image stored in the database. If the image is not aligned properly then the concept of window size comes into picture. According to (5), the number of calculations increases by a factor of $(WindowSize)^2$. The complexity calculation algorithm of SWM is explained below.

$$ComplexitySWM = M \times N \times (WS)^2 \times O \times T_1 \tag{8}$$

where, $M = (No.ofRows - WS) = (60 - WS)$, $N = (No.ofColumns - WS) = (60 - WS)$

$WS = WindowSise = 1to4$, $O = No.ofOrientations = 7$, $T_1 = EX - ORTime$

According to (5), if $WS = 4$ and time taken for each EX-OR operation is T_1 as shown in (9), then total time taken for hamming distance calculation is given by:

$$(FV(i, j) \oplus FV_{DB}(i, j)) = T_1 \text{ time} \tag{9}$$

Complexity or total sliding window time (T_{SWM}) taken for 6 orientations will be given by (10):

$$T_{SWM} = (60 - 4) \times (60 - 4) \times (4 \times 4) \times 7 \times T_1 = 56 \times 56 \times 16 \times 7 \times T_1 = 351232T_1 \tag{10}$$

where, T_{SWM} specifies time taken to compare feature vectors of two palmprints.

If we consider 100 palmprints the total matching time will be given by:

$$((351232T_1 \times 100) = 35123200T_1)$$

It is observed that the number of operations is large in number in sliding window method and will take lot of time. Comparison time need to be reduced by improving sliding window method.

4.2 Sliding Window Method 1 (SWM1)

In this method, a part of the actual palmprint area is considered and it is named as palmprint segment. The chosen palmprint area can be any of the palmprint segment mentioned in the Fig. 7.

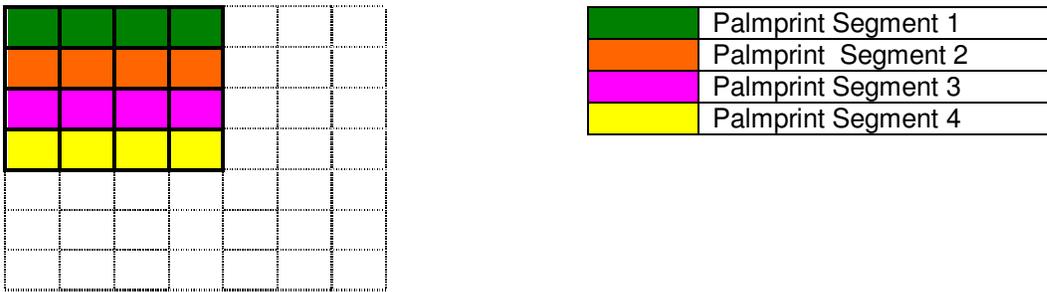


FIGURE 7: The segmented palmprint.

The complexity calculation algorithm of SWM1 is explained below.

$$ComplexitySWM1 = (M_1 \times N \times (WS)^2 \times T_1 + M \times N \times T_1) \times O \tag{11}$$

$$\text{where, } M_1 = \left(\frac{No.ofRows}{4} - WS \right) = \left(\frac{60}{4} - WS \right) = (15 - WS),$$

$$M = (No.ofRows - WS) = (60 - WS), N = (No.ofColumns - WS) = (60 - WS)$$

$WS = WindowSise = 1to4$, $O = No.ofOrientations = 7$, $T_1 = EX - ORTime$

In (6) minimum hamming distance value using sliding window method is calculated. In SWM1 method, hamming distance value for o orientations is found out using sliding window method as shown in (5). The (6) signifies minimum hamming distance in sliding window as the closest matching between two palmprints.

$$\min_index_HD_{WS} = index(\min(HD_{WS})) \tag{12}$$

Total time for 6 orientations and window size 4 will be given by (11):

$$T_{SWM1} = (((11 \times 56 \times 16 \times T_1 + 56 \times 56 T_1) \times 7) = 90944T_1).$$

If we consider 100 palmprints the total matching time will be $((90944T_1 \times 100) = 9094400T_1)$.

The reduction in matching time is given by (13).

$$\frac{(351232T_1 - 90944T_1)}{351232T_1} \times 100 = 74.11\% \tag{13}$$

There is 74.11% reduction in matching time using SWM1 than SWM. The reduced matching time signifies fast authentication system.

The sliding window method on the chosen palmprint area can be shown diagrammatically as in Fig. 8.

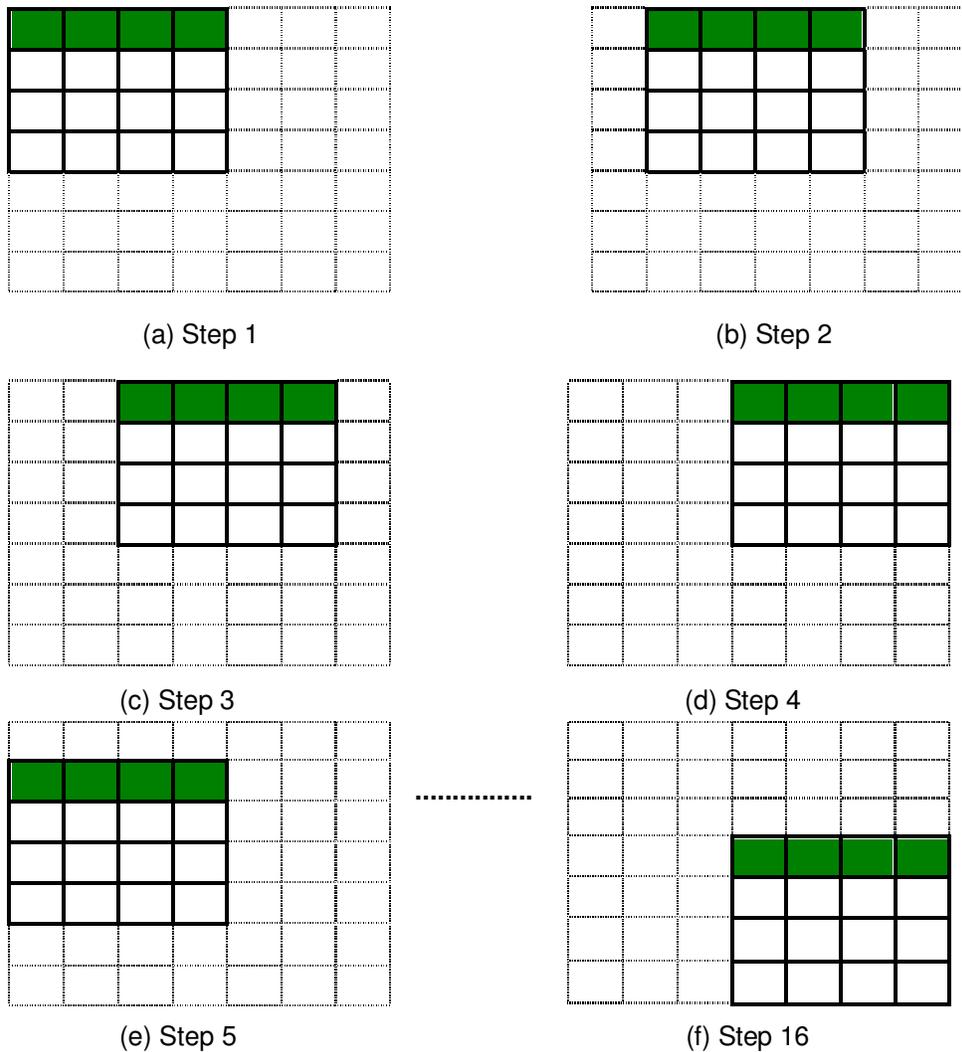


FIGURE 8: Various steps in improved Sliding window method

As we can see from the above diagrams that the area of palmprint segment in sliding window method has reduced considerably and it leads to reduction in matching time.

4.3 Sliding Window Method 2 (SWM2)

In this method, the minimum hamming distance value is not calculated for all the orientations using sliding window method. The index of minimum hamming distance value is calculated for one orientation and same index value will be used to calculate hamming distance for other orientations. The complexity calculation algorithm of SWM2 is explained below.

$$Complexity_{SWM2} = \left((M_1 \times N \times (WS)^2 \times M \times N) + (M \times N) \times (O - 1) \right) \times T_1 \quad (14)$$

where, $M_1 = \left(\frac{\text{No.ofRows}}{4} - WS \right) = \left(\frac{60}{4} - WS \right) = (15 - WS),$

$M = (\text{No.ofRows} - WS) = (60 - WS), N = (\text{No.ofColumns} - WS) = (60 - WS)$

$WS = \text{WindowSise} = 1\text{to}4, O = \text{No.ofOrientations} = 7, T_1 = \text{EX} - \text{ORTime}$

The total time taken given by (14) for window size 4, $T_{SWM2} = ((16 \times 11 \times 56 \times T_1) + (56 \times 56 \times T_1) \times 6) = 31808T_1.$ If we consider 100 palmprints the total matching time will be $(31808T_1 \times 100 = 3180800T_1).$

$$\frac{(351232T_1 - 31808T_1)}{351232T_1} \times 100 = 90.94\% \tag{15}$$

There is 90.94% reduction in matching time using SWM2 than SWM. The reduced matching time signifies fast authentication system.

All the sliding window methods are summarized in table 1.

Method used	Number of operations
1. SWM: Hamming distance calculation with sliding window method with window size $WS.$ $HD_{WS} = \sum_i^{60-WS} \sum_j^{60-WS} (FV(i, j) \oplus FV_{DB}(i, j))$	$56 \times 56 \times 16 \times 7 = 351232 \text{ Operations}$ The (56×56) implies the number of comparisons of the palmprint with the palmprint in the database. $(56 \times 56 \times 16)$ signifies sliding window method comparisons for window size 4, so $(4 \times 4) = 16$ comparisons. The total number of comparisons takes place for seven orientations is 351232.
2. SWM1: Hamming distance calculation with sliding window method (15×60) with window size 4. $HD = \sum_i^{15-WS} \sum_j^{60-WS} (FV(i, j) \oplus FV_{DB}(i, j))$ The palmprint segment size is $(15 \times 60).$	$(11 \times 56 \times 16 + 56 \times 56) \times 7 = 90944 \text{ Operations}$ The (11×56) implies the number of comparisons of the palmprint segment with the palmprint segment in the database. $(11 \times 56 \times 16)$ signifies to sliding window method comparisons for window size 4, so $(4 \times 4) = 16$ comparisons. (56×56) signifies the comparison of palmprint at the minimum index value. The sum of $(11 \times 56 \times 16)$ and (56×56) give the number of comparison for each orientation. The total number of comparisons takes place for seven orientations is 90944 Operations.
Assumption: If time taken to do each matching operation is $T_1.$	Theoretical time improvement $\frac{(351232T_1 - 90944T_1)}{351232T_1} \times 100 = 74.11\%$
3. SWM2: Hamming distance calculation with sliding window method (15×60) with window size $WS.$ $HD = \sum_i^{15-WS} \sum_j^{60-WS} (FV(i, j) \oplus FV_{DB}(i, j))$	$(11 \times 56 \times 16 + 56 \times 56) + (56 \times 56) \times 6 = 31808$ Operations The assumption here is that the minimum hamming distance values for other orientations will also be at the same index as it is for one orientation. $(11 \times 56 \times 16)$ applies to sliding window method comparisons for window size 4, so

	<p>$(4 \times 4 = 16)$ comparisons. (56×56) signifies the comparison of palmprint and finding the minimum index value. The sum of $(11 \times 56 \times 16)$ and (56×56) give the number of operations for one orientation. The number of comparisons for another 6 orientation is at the minimum index value as assumed for this method. So, additional $(56 \times 56) \times 6$ comparisons will be added. The total number of comparisons is 31808.</p>
Assumption: If time taken to do each matching operation is T_1 .	<p>Theoretical time improvement</p> $\frac{(351232T_1 - 31808T_1)}{351232T_1} \times 100 = 90.94\%$

TABLE 1: Reduction in matching operations using various Sliding window method

The number of operations, matching operation time with respect to sliding window size is tabulated in table 2.

Sliding Window Size	Number of Operations			Matching operation Time			Percentage reduction in Matching time		
	SWM	SWM 1	SWM2	SWM	SWM1	SWM2	SWM	SWM1	SWM2
1	24367	24367	24367	4.48E-03	4.45E-03	4.43E-03	NA	0.67	1.12
2	94192	44660	26564	1.75E-02	8.34E-03	4.98E-03	NA	52.34	71.54
3	204687	65835	28899	3.89E-02	1.23E-02	5.53E-03	NA	68.38	85.78
4	351232	90944	31808	6.53E-02	1.69E-02	6.32E-03	NA	74.12	90.32

TABLE 2: Percentage reduction in matching operation time

From table 2 it is concluded that there is maximum of 74.12% and 90.32% practical reduction in matching operation time. So it can be concluded from table 1 and table 2 results that with two different sliding window methods, the authentication system becomes very fast. Figure 9 shows that with SWM1 and SWM2 the comparison time reduces significantly.

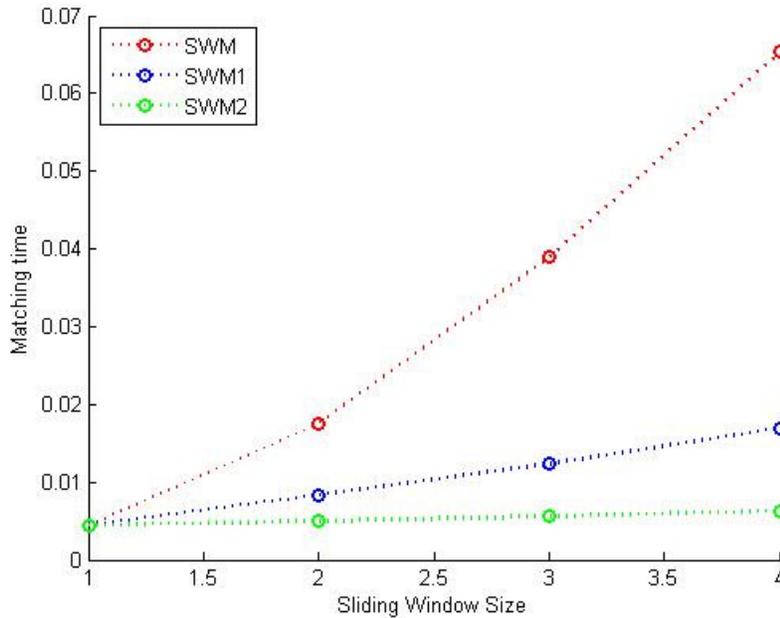


FIGURE 9: Matching time Vs Sliding Window size

5. REFERENCE THRESHOLD

The hand image samples are divided into two groups G1 and G2.

G₁ group

$$P_1 = [I_1, I_2, \dots, I_{(M-1)}], P_2 = [I_1, I_2, \dots, I_{(M-1)}], \dots, P_N = [I_1, I_2, \dots, I_{(M-1)}] \tag{16}$$

G₂ group

$$P_1 = [I_M], P_2 = [I_M], \dots, P_N = [I_M] \tag{17}$$

where P_i denotes ith person in group G₁, G₂, I_j denotes the jth palm image in group G₁, G₂.

i \ j	1	2	3	M-1
1	X	HD ₁₂	HD ₁₃	HD _{1(M-1)}
2	HD ₂₁	X	HD ₂₃	HD _{2(M-1)}
⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮
M-1	HD _{(M-1)1}	HD _{(M-1)2}	HD _{(M-1)3}	X

TABLE 3: Matching in group G₁ among person P₁

In group G₁, each hand feature vector in P₁ is matched with all other (m-1) hands feature vector by Hamming Distance similarity measurement method. The matching values are approaching “0” and are stored in threshold array.

$$TA_1 = [HD_{12}, HD_{13}, \dots, HD_{1(M-1)}, HD_{21}, HD_{23}, \dots, HD_{2(M-1)}, \dots, HD_{(M-1)1}, HD_{(M-1)2}, \dots, HD_{(M-1)(M-2)}] \tag{18}$$

Similarly, all N hand image samples matching results are stored in Threshold array (TA).

$$T_A = TA_1 + TA_2 + \dots + TA_N \tag{19}$$

The minimum and maximum of matching values are found out from the threshold array (TA₁, TA₂, TA_N) for each individual as shown in equation (20).

$$\left. \begin{aligned} T_{AiMIN} &= \min(T_{Ai}) \\ T_{AiMAX} &= \max(T_{Ai}) \end{aligned} \right\}_{i=1, \dots, N} \quad (20)$$

The accuracy of the system is identified by matching group G_2 samples with group G_1 samples using threshold values stored in threshold array. Finally, a threshold value is chosen where FAR and FRR is minimum, this value is called Reference threshold.

6. EXPERIMENTAL RESULTS

The effectiveness of the proposed Complex Gabor Wavelet method for feature extraction technique was assessed on the PolyU database [16]. The database consists of 600 palmprint images from 100 palms with 6 samples for each palm.

6.1 Palmprint Authentication System

For the verification experiments the database has to be partitioned into two non-overlapping groups. The first group (G_1) represents the training group, consisting of 100 persons with each person having 5 palm sample images whereas the second group (G_2) represents the testing group containing 100 persons with each person having one palm image different from the first group images. The image size is 284×384 pixels.

G_1 group

$$P_1 = [I_1, I_2, I_3, I_4, I_5], \quad P_2 = [I_1, I_2, I_3, I_4, I_5], \dots, \quad P_{100} = [I_1, I_2, I_3, I_4, I_5]$$

In G_1 group each hand P_i contains 5 sample image I_{1-5} .

G_2 group

$$P_1 = [I_6], P_2 = [I_6] \dots \dots \dots P_{100} = [I_6]$$

In G_2 group each hand P_i contains only sample image I_6 .

Image is pre-processed to get the region of interest. Pre-processing includes image enhancement, image binarization, boundary extraction, cropping of palmprint/ROI. The ROI size is 60×60 pixels. Feature extraction is done by Complex Gabor Wavelet Transform to get the imaginary wavelet coefficients. The Complex Gabor coefficients are calculated at seven different orientations respectively. The feature vector contains features extracted by Complex Gabor wavelet at different orientations. Feature vector of all hand images samples is calculated and stored in database. Feature vector matrix is $CGWF = [CGWF_0, CGWF_1, CGWF_2, CGWF_3, CGWF_4, CGWF_5, CGWF_6]$. Sample of ROI is shown in Fig.10.

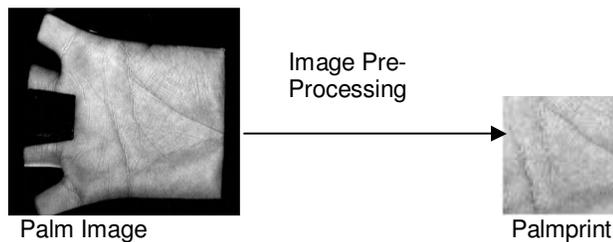


FIGURE 10: Sample of ROI.

Hamming distance is used as a similarity measurement method for feature matching.

6.2 Reference Threshold Calculation

In group G_1 , each hand feature vector in P_i is matched with all other 4 hands feature vector by Hamming Distance similarity measurement method. Similarly, all 100 hand image samples 2000 matching values are stored in Threshold array (TA).

$$T_A = TA_1 + TA_2 + \dots + TA_{100}$$

$$\left. \begin{aligned} T_{AiMIN} &= \min(T_{Ai}) \\ T_{AiMAX} &= \max(T_{Ai}) \end{aligned} \right\}_{i=1, \dots, 100}$$

The maximum and minimum values are found out from threshold array.

These 25 threshold values are tested with group G_2 and group G_1 images. The 25 threshold values, the respective FAR and FRR values and accuracy at frequency sixteen and theta zero are shown in table 4.

Reference Threshold	FAR	FRR	Accuracy
5.46E-01	1.23E-01	6.69E-03	9.35E+01
5.52E-01	1.27E-01	4.91E-03	9.34E+01
5.58E-01	1.16E-01	2.67E-03	9.40E+01
5.65E-01	9.25E-02	1.33E-03	9.53E+01
5.71E-01	6.95E-02	6.94E-03	9.62E+01
5.77E-01	4.92E-02	3.74E-04	9.75E+01
5.84E-01	3.59E-02	2.75E-04	9.82E+01
5.90E-01	2.71E-02	3.13E-04	9.86E+01
5.96E-01	2.16E-02	1.85E-04	9.89E+01

TABLE 4: Threshold Values, FAR, FRR, Accuracy Values for SWM.

Method	FAR	FRR	Accuracy
SWM	6.95E-02	6.94E-03	9.62E+01
SWM1	6.94E-02	7.13E-03	9.62E+01
SWM2	7.07E-02	6.93E-03	9.61E+01

TABLE 5: FAR, FRR, Accuracy values for sliding window size 2

From the table 5 it can be concluded that by using SWM1 and SWM2 techniques the accuracy of the system is not affected.

The FAR and FRR are values are plotted with respect to 25 threshold range values. From the graph the value of reference threshold is chosen where FAR and FRR are minimum. Plot of FAR and FRR is shown in figure 11. The plot between accuracy and threshold is shown in figure 11.

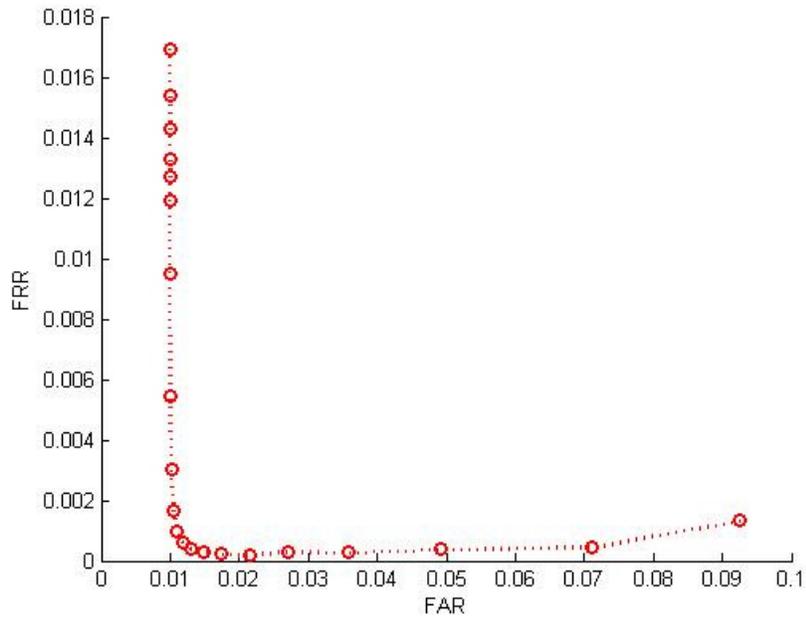


FIGURE 11: FAR Vs FRR

Accuracy values with respect to threshold values are plotted in Fig. 12.

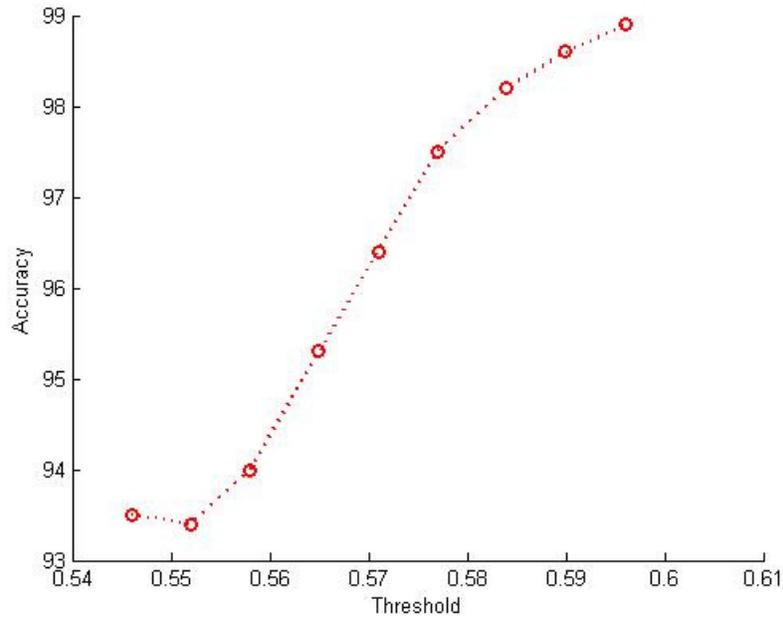


FIGURE 12: Accuracy Vs Threshold

6.3 Speed Performance

Table 1 shows the reduction in comparison time by sliding window method 1 (SWM1) and sliding window method 2 (SWM2). It can be observed that with palmprint segments, the comparison time reduces and speed to verify the person is improved.

7. CONCLUSION

The Complex Gabor Wavelet transform is applied to each palmprint image in the database to get imaginary wavelet coefficients. The features are computed at different orientations. The feature vector generated from complex gabor features are matched by hamming distance using sliding window method. Modified sliding window methods SWM1 and SWM2 are also used to make the matching system fast. Experimental results clearly show that our proposed Complex Gabor Wavelet method can discriminate palmprints and is fast authentication system.

8. REFERENCES

- [1] Jain A.K., Ross A., Prabhakar S.: 'An introduction to biometric recognition', IEEE Trans. Circuits Syst. Video Technol., 14, (1), pp. 4–20, 2004.
- [2] Pavešlc´ N., Ribarić S., Ribarić D.: 'Personal authentication using hand-geometry and palmprint features – the state of the art'. Proc. Workshop: Biometrics – Challenges Arising from Theory to Practice, Cambridge, pp. 17–26, 2004.
- [3] Kumar and D. Zhang, "Combining fingerprint, palmprint and handshape for user authentication," In *Proceedings of ICPR*, vol.4, pp.549- 552.
- [4] Kumar A., Zhang D.: 'Personal authentication using multiple palmprint representation', Pattern Recognit., 38, (10), pp. 1695–1704, 2005.
- [5] Zhang, D., Kongi, W., You, J., and Wong, M.: "Online palmprint identification", IEEE Trans. Pattern Anal. Mach. Intell., vol. 25, no. 9, pp. 1041–1049, 2003.
- [6] B.S. Manjunath, "Gabor Wavelet Transform and Application to Problems in Early vision", IEEE, pp. 796-800.
- [7] Wu, X., Zhang, D., and Wang, K.: "Fisherpalms based palmprint recognition", Pattern Recognit. Letters, vol. 24, no. 15, pp. 2829–2838, 2003.
- [8] Hu, D., Feng, G., and Zhou, Z.: "Two dimensional locality preserving projections with its applications to palmprint recognition", Pattern Recognit, 40, pp. 339–402, 2007.
- [9] Lu, G., Zhang, D., and Wang, K.: "Palmprint recognition using eigenpalms features", Pattern Recognit. Lett., vol. 24, no.9–10, pp. 1463–1467, 2003.
- [10] Liu, C.: "Gabor-based kernel PCA with fractional power polynomial models for face recognition", IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, no. 5, pp. 572–581, 2004.
- [11] J. Daugman, "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression," IEEE Trans. Acousr. Speech Signal Proc., vol. 36, no. 7, pp. 1169-1 179, 1988.
- [12] Sylvain Fischer and Gabriel Cristdhal, "Minimum Entropy Transform Using Gabor Wavelets For Image Compression", IEEE, pp 428-433, 2001.
- [13] Jain and F. Farrokhnia, "Unsupervised texture segmentation using gabor filters. Pattern Recognition, 24(12):1167– 1186, 1991.
- [14] Say Song Goh, Amos Ron, Zuowei Shen "Gabor and Wavelet Frames".

- [15] Muwei Jian, Haoyan GUO, Lei Liu, "Texture Image Classification Using Visual Perceptual Texture Features and Gabor Wavelet Features", Journal Of Computers, Vol. 4, No. 8, Academy Publisher, pp 763-770, Aug. 2009.
- [16] The PolyU Palmprint Database: <http://www4.comp.polyu.edu.hk/biometrics/>

Snow Cover Estimation from Resourcesat-1 AWiFS – Image Processing With an Automated Approach

S. Subramaniam

subramaniam_s@nrsc.gov.in

*RS & GIS Applications Area
National Remote Sensing Centre
Indian Space Research Organization (ISRO)
Department of Space Govt. of India
Balanagar, Hyderabad – 500625
INDIA*

A.V. Suresh Babu

avsureshbabuisro@ gmail.com

*RS & GIS Applications Area
National Remote Sensing Centre
Indian Space Research Organization (ISRO)
Department of Space Govt. of India
Balanagar, Hyderabad – 500625
INDIA*

E. Sivasankar

sivasankar_e@nrsc.gov.in

*RS & GIS Applications Area
National Remote Sensing Centre
Indian Space Research Organization (ISRO)
Department of Space Govt. of India
Balanagar, Hyderabad – 500625
INDIA*

V. Venkateshwar Rao

vvrao@nrsc.gov.in

*RS & GIS Applications Area
National Remote Sensing Centre
Indian Space Research Organization (ISRO)
Department of Space Govt. of India
Balanagar, Hyderabad – 500625
INDIA*

G. Behera

behera_g@nrsc.gov.in

*RS & GIS Applications Area
National Remote Sensing Centre
Indian Space Research Organization (ISRO)
Department of Space Govt. of India
Balanagar, Hyderabad – 500625
INDIA*

Abstract

Snow and glaciers cover large areas of Himalayas. The resulting runoff from snow and glacier melt provides nearly 30-50% of the total annual water outlay of most of the rivers in north India. There is a need for continuous monitoring of the Himalayan snow cover area at more frequency and over longer period in the context of climate change studies. Normalized Difference Snow index (NDSI) technique which generally used for automated detection of snow cover from remotely sensed data has limitations in the detection of snow under shadow and exclusion of water and gray cloud. A new automated snow cover estimation algorithm has been developed and presented here to overcome the these limitations using the spectral information available in the Red, Near Infrared and SWIR spectral bands of IRS P6 Resourcesat-1 AWiFS sensor. The automated algorithm has been implemented in hierarchical logical steps. Algorithm has been validated by comparing the results obtained with Hall's and Kulkarni's methods and observed that

the new algorithm performs better than other methods in the elimination of noise like glacial water bodies and water bearing dark clouds, while detecting the snow covered pixels in deep mountain shadows. Satisfactory results have been obtained, when used with several temporal images of large image mosaics covering different regions of Himalayas from Kashmir in the west to Arunachal Pradesh in the East covering entire snow belt of Indian region. This has shown the robustness of the algorithm in various locations. Intra annual and inter annual snow cover over Himalayas regions of the India was evaluated and the results were encouraging and are suitable for porting on to a public domain. This algorithm has been evaluated with Landsat ETM and IRS LISS III which has similar spectral bands with different spatial and radiometric resolutions and the algorithm has been found to be working satisfactorily. The algorithm has been found to be useful for regular periodic monitoring of snow cover area and is generic in nature that can be used with various sensor data that has similar spectral bands.

Keywords: Snow cover , Automated algorithm, Remote Sensing.

1. INTRODUCTION

Snow and glaciers cover large areas in Himalayas during winter. The spatial extent of snow undergoes significant change with drop in ambient temperature during winter and snow melt with the commencement of spring. During spring and summer, such a change impacts the stream flow of the rivers originating in the higher Himalayas which are fed by snow and glacial melt run off. Recent study carried out on Himalayan glaciers [1, 2] and National Institute of Natural Resources on the mountain glaciers, Peru [3] reveals that the glaciers are receding rapidly due to the global warming and climate change. Recent reports on climate change due to global warming has predicted rapid change in the snow line and glaciers in Himalayas and its impact on the food production in the South Asian countries. Hence, regular monitoring and mapping of snow cover area is necessary for water resources planning and management in snow fed river basins for the assessment of water resources. The remote sensing technique has been used extensively for snow cover monitoring [4-16] and snow melt run off studies [7] for Himalayan glaciers. Various analysis techniques such as visual, hybrid visual-digital and supervised Maximum Likelihood Classification, have been used to estimate the aerial extent of snow cover.

In the context of global warming and the resulting climate change scenario, there is a need for near real time monitoring of dynamic natural resources like snow cover which involves handling and analysis of large volume satellite data and carry out simulation studies related to climate change or impact of climate change. The availability of large quantity of data from high-resolution space-based sensors along with the existing archives surpassed our ability to extract information from all the data using supervised image analysis techniques. The supervised image information extraction process is complex, time consuming and expensive in nature. This hinders the use of all acquired and accessible data, penalizing large environmental monitoring projects, and might even leave critical phenomena totally undetected.

Presently, earth observation (EO) data are archived and disseminated based on the metadata such as geographical location, date of acquisition and sensor type, which provide no information on the information content of the data. Experts interpret these images to extract information using their domain knowledge and the service provider/user combines them with other legacy information to support decision-making. In this scenario, the offer of data does not match the customer's real need, which is image information. In response to these needs, attempts are being made to develop automatic classifiers [18-20] suitable for the extraction of predefined and specific type of information from the given satellite sensor data with improved throughput in image analysis for better utilization of satellite data by the user. Survey of various classifiers for the analysis of remote sensing data and their comparative analysis showed the knowledge based classifiers provided improved accuracies over others [21].

Two types of automated classifiers are being defined and used for the analysis of remotely sensed data. Walter(1998) has used GIS database of LULC map for the generation of training sample, for automatic classification and updating of GIS database. In this method, the training sample data sets for all the classes are automatically generated using the information available in the GIS database and then the classification is carried as in the supervised classifier. This method assumes that the wrongly captured GIS objects and the number of changes in the real world are substantially less than all the number of GIS objects of the data set. The usefulness of knowledge base, used in these classifiers, is limited to the image data set used for extracting the training samples. The other method is a knowledge-based method [22,23], which relies on invariant / semi invariant feature parameter corresponding to different land use/land cover feature categories and it is most suitable for automatic classification. The spectral reflectance, being the invariant property of objects, has to be used for the extraction of the feature parameters.

Hall et al., [6-7] proposed the use of Normalized Difference Snow Index (NDSI) along with NIR reflectance threshold for snow cover mapping. It has been found that the NDSI based method provide better accuracy in the classification of results compared to the supervised Maximum likelihood classifier, due to the difficulties encountered in defining pixels under shadow that were snow covered without inadvertently mapping non-snow pixels [6]. However, the major difficulty in snow cover monitoring using NDSI based automated technique is elimination of snow under mountain shadow [6-7], inclusion of glacial water bodies ([11]) and gray clouds. These rain bearing clouds occur in the monsoon season. Hence, the combination of digital and visual interpretation techniques has been used for snow cover mapping, which is time consuming. An automated extraction algorithm for the identification of snow cover area using NDSI method has been developed to overcome the above limitations using the information existing in all the four bands of Resourcesat -1 AWiFS sensor. The importance of the usefulness of the Red, NIR, SWIR spectral bands in the elimination of water in the presence of the snow covered pixels under shadow and grey cloud has been brought out. The methodology for the algorithm development, validation of the results on large image mosaics has been reported in this paper

2. METHODOLOGY

Automated image analysis technique for extraction of information from the remote sensing image data must be scene (geographic location) and season independent. The spectral reflectance pattern is the intrinsic property of an object and the objects belonging to a given land Use / land cover class has similar shape, even though it can have variation in the magnitude due the intrinsic variations of the class features because of weathering, wetness etc., and also due to the extrinsic variations due to topography and sun illumination. The discriminating feature parameters that represent the spectral shape such as normalized spectral indices, spectral band ratios and brightness values being used in the classification of the remotely sensed image data. These parameters essentially capture the contrast features present between the feature classes and with-in class feature contrasts. The with-in class feature contrast captures the characteristic feature of the particular class, for example : the large reflectance change between the visible and short wave infrared band for the snow/ice. The between class contrast captures the discriminating features between the given set of object classes. The with-in class contrast features have been used for identification of the feature class, while the between class contrast are used for the elimination of other feature classes in the automated delineation of the specific feature classes such as water, snow, urban areas.

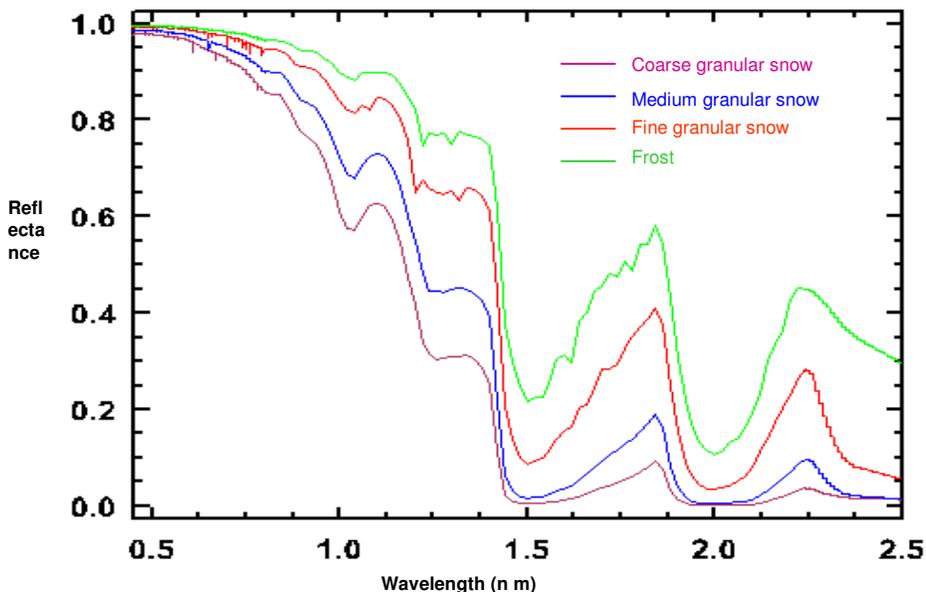
Development of automatic extraction algorithm involves the collection and analysis of spectral information (knowledge base), modeling their relationships, development and implementation of hierarchical logical algorithm, implementation of algorithm on various seasonal datasets, validation of the algorithm by comparison with standard NSDI based methods and hybrid digital visual interpretation techniques. The knowledge based method has been implemented in multi-logic decision equation using spectral spatial and contextual based parameters with suitable threshold levels. The selected parameter either support and reinforce the identification of the particular feature class or eliminate the noise feature classes. It is possible to identify and extract the pixels corresponding to the various feature categories with sufficient number of logical

decisions. A separate logical equations are used for subcategory of the feature classes, to account for the natural variations in the feature classes and also for the variations due to the geo-climatic impact on the spectral response pattern of the feature class. The threshold values used in these logics are flexible as these are used to reinforce the decision taken with multiple logics and is similar to the approach used by human interpreter in analyzing the image data in visual interpretation technique. The logical equations in general contain one or two essential conditions to identify the presence of a particular class and the other conditions are used for the removal of the noisy feature classes that may have similar response in certain geo-climatic conditions.

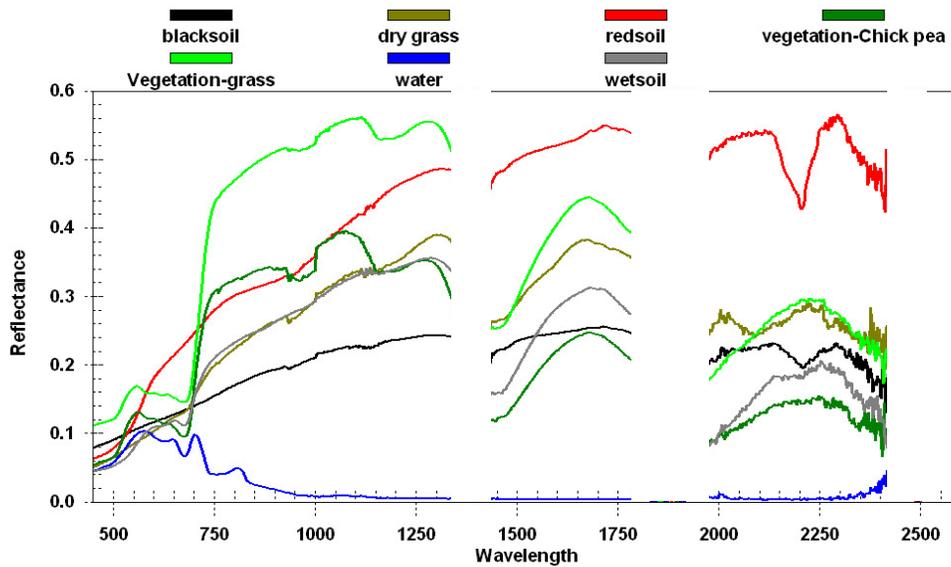
In the case of snow covered pixels / area extraction, spectral information only has to be used and the image information of the area such as texture is not useful due the reason that texture does not represent underground or the presence or absence of the snow cover over the ground. Hence, in snow cover extraction methods, the spectral feature parameters that represent the shape of the reflectance pattern of the snow and discriminating characteristics with respect to the other land cover classes and noise classes like cloud and cloud shadow also has to be used. The normalized difference spectral indices such as NDSI, NIR-SWIR indices and spectral ratios have been computed from the multi spectral image data and for the extraction of the snow covered pixels in the algorithm.

2.1 Analysis of Spectral Reflectance Pattern of Snow

The field measurements of spectral response pattern of different types of snow and ice has been carried out [4, 11] and also in the laboratory at John Hopkins university (spectral library of snow types) and the laboratory measurements is shown in figure 1a.. The spectral reflectance of snow has high values in visible bands and has small value in short-wave infrared band (1550-1750 nm). A comparison of the spectral response pattern of snow/ice cover (figure 1a) and other land cover types (figure. 1b) reveals that (i) snow has higher spectral contrast between the Green and SWIR spectral region compared to others except water ; (ii) snow has relatively higher reflectance in the visible region compared to that of other land types like water, soil, vegetation ; and (iii) Snow has higher reflectance values similar to cloud in visible region, while in the SWIR region it has much lower reflectance and this has been used as discriminating parameter.



(a) Snow / ice



(b) land cover types

FIGURE 1: Spectral reflectance characteristic of (a) snow/ice and (b) different land cover types.

In case of air or space borne measurements, the image DN values representing the satellite radiance is proportional to the ground object and spectral reflectance response pattern has been modified due to scattering and absorption by atmospheric constituents, variations of sun's illumination geometry, and sensor, etc. The conversion of image DN values to Top Of the Atmosphere (TOA) reflectance value [24] has been used to account for illumination of sensor geometry except for the atmospheric effect. Hence, TOA reflectance values for development of the automated snow cover extraction algorithm. The TOA reflectance (pTOA) is computed using the following equation.

$$\rho_{TOA} = \left(\frac{\pi L_{\lambda} d^2}{ESUN \cos(\Theta_s)} \right) \text{-----(1)}$$

- where ρ_{TOA} = TOA reflectance;
- L_{λ} = spectral radiance at sensor aperture ($mW \text{ cm}^{-2} \text{ sr}^{-1} \mu^{-1}$);
- d = Earth-Sun distance (Astronomical Units);
- $ESUN$ = mean solar exo-atmospheric irradiance ($mW \text{ cm}^{-2} \text{ sr}^{-1} \mu^{-1}$);
- Θ_s = solar zenith angle (degrees).

2.2 Normalized Difference Snow Index (NDSI)

Hall et.al [7] proposed the Normalized Difference Snow Index (NDSI) for the detection of pixel having snow and ice cover types based on analysis of the spectral reflectance pattern.

$$NDSI = \frac{\text{Green reflectance} - \text{SWIR reflectance}}{\text{Green reflectance} + \text{SWIR reflectance}} \text{----- (2)}$$

A scatter plot (Figure.2) of NDSI as a function of TOA reflectance in NIR Spectral band for various land use / land cover features indicates that riverine sand/ barren rock, urban and vegetation have negative NDSI values and cloud has value close to 0.2. The NDSI value for

snow/ice has been found to be above 0.4, while water has values in the range of 0.2 to 0.6. The shallow water and mixed pixels of water and soil / sand has lower NDSI values. Sensitivity analysis has been carried out by Kulkarni et al.2006 to know the effect of NDSI threshold on the snow cover area estimation indicated a drop in area of approximately 3% for NDSI threshold of 0.4, beyond which the reduction of snow cover area increases rapidly with higher NDSI threshold and concluded that the NDSI value of 0.4, provides reasonable accuracy for snow cover monitoring. Hall et al [7] has proposed a threshold value of $NDSI > 0.4$ and NIR reflectance > 0.11 for snow cover estimation using TM spectral bands. The same threshold values were also used by Xiato et al [15, 16] with SPOT4 Vegetation (VGT) sensor. The NIR reflectance threshold of 0.11 has been used to eliminate water pixels in the snow cover estimation which lead to exclusion of snow under mountain shadows. Kulkarni et al [11] used a threshold value of $NDSI > 0.4$ to include the snow under shadow and suggested pre-enumerated water body mask for eliminating the water in snow estimation. The spectral reflectance measurements [10,11] has shown that the NDSI is suitable for the detection of the snow cover even under mountain shadow. In view of above, the possibility of using information from other spectral bands to overcome the above limitation has to be studied.

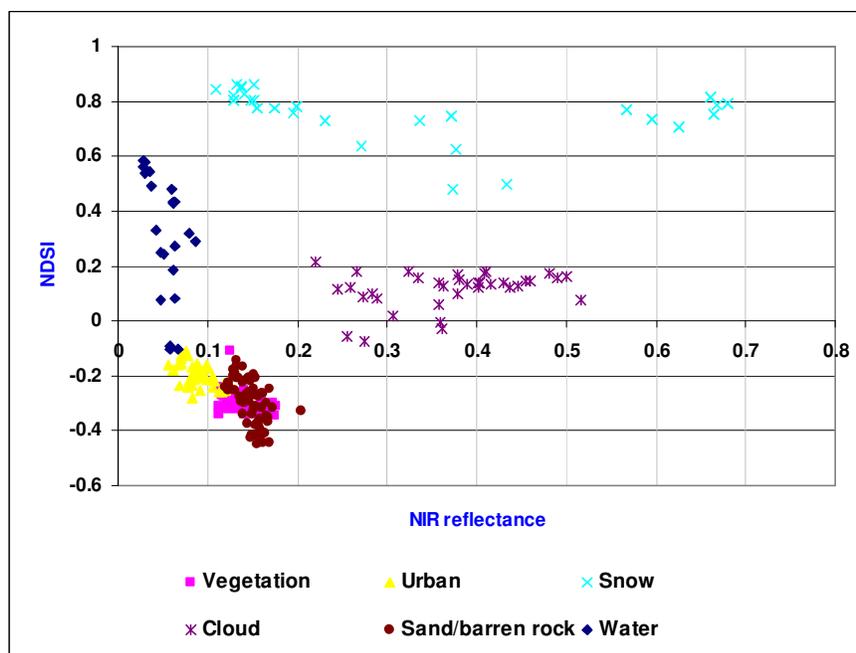


FIGURE 2 : TOA reflectance derived NDSI scatter plot for snow and other land cover types.

2.3. Spectral Reflectance Pattern Analysis and Algorithm Development for Snow Cover Estimation

The snow covered pixels of various levels of illumination (directly illuminated by SUN and those under different levels of shadow cover and at various seasons) were identified by visual observations of the TOA image and their TOA values were extracted. Further, the TOA reflectance values of the noise features such as water and grey cloud pixels were extracted after identifying them visually from different images . Analysis of the spectral reflectance pattern of different snow under different illumination conditions based on visual observations and noise features such as water, grey cloud has been carried out. The scatter plot of NDSI as a function of NIR reflectance for different types of objects including snow under different illumination conditions (that can have NDSI value greater than 0.4) are shown in Figure 3. The NIR reflectance > 0.11 for eliminating water bodies, leads to the exclusion of snow pixels under shadow. If this NIR threshold has been removed, then different method has to be used for the elimination of water pixels for snow cover estimation. Kulkarni et al [11] have suggested the use of water body mask

to remove the water pixels being classified as snow. This may not be suitable for all geographic locations and seasons. To solve the problem of water pixels, a comparative analysis of the spectral characteristics water and snow under shadow has been carried out to identify discriminating features.

2.3.1. Elimination of Water from Snow

The analysis has shown that there exists differences in the spectral reflectance of water (figure.4a), snow under shadow (figure.4b) in the Red–NIR spectral region which can be used for their discrimination. The scatter plot of NDSI as a function of normalized NIR-SWIR difference index (NIR-SWIR / NIR+SWIR) and Red-NIR ratio (Red/NIR) for various types of snow cover and water has been shown in figure 5. It has been observed that a combination of NIR- SWIR index value and Red-NIR ratio can provide the required discrimination of snow from water which has been implemented with suitable threshold along with NDSI in the development of algorithm.

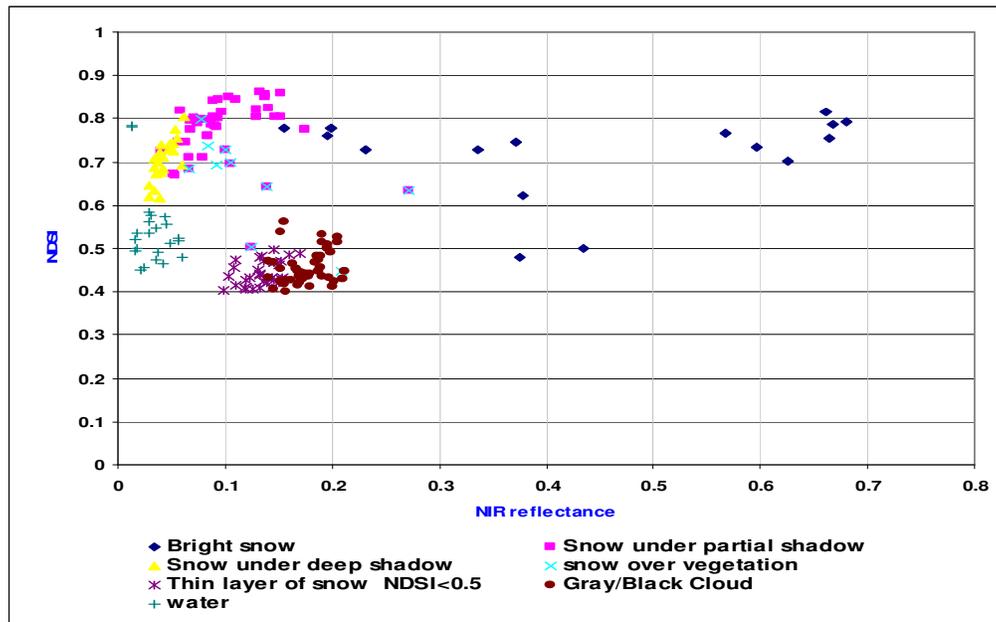
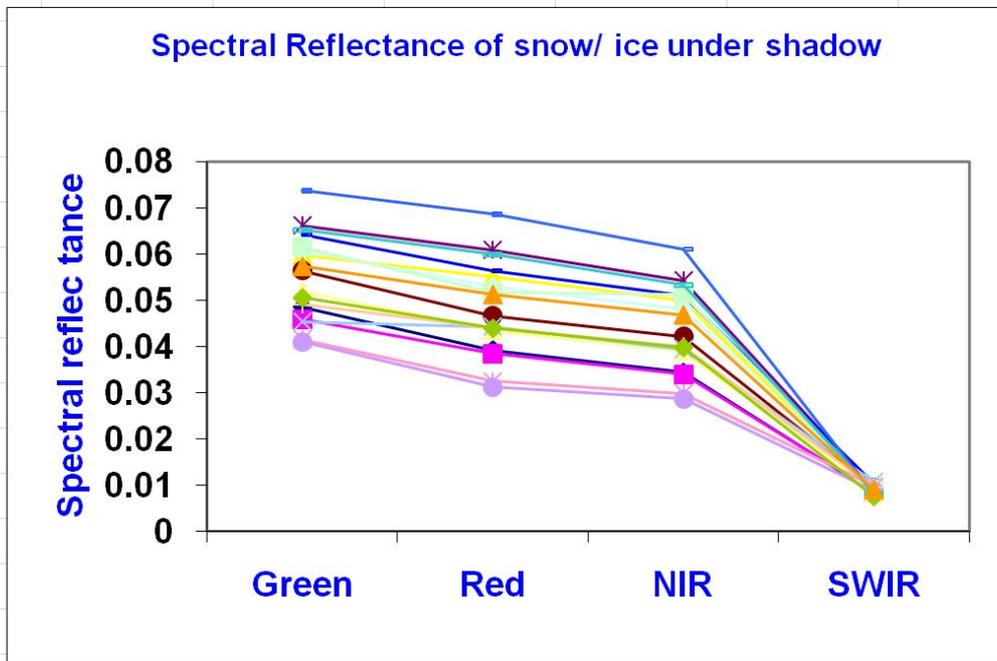
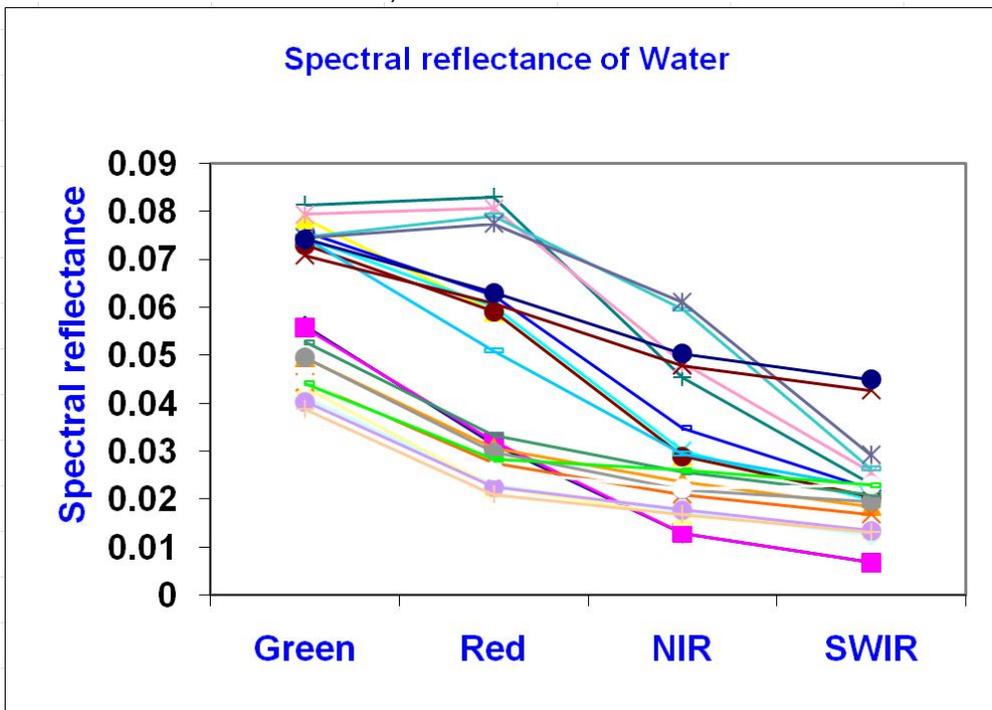


FIGURE 3: Scatter plot of normalized difference snow index (NDSI) as function of NIR reflectance.



a) snow under shadow



(b) water

FIGURE 4: Spectral reflectance pattern of water and snow under shadow.

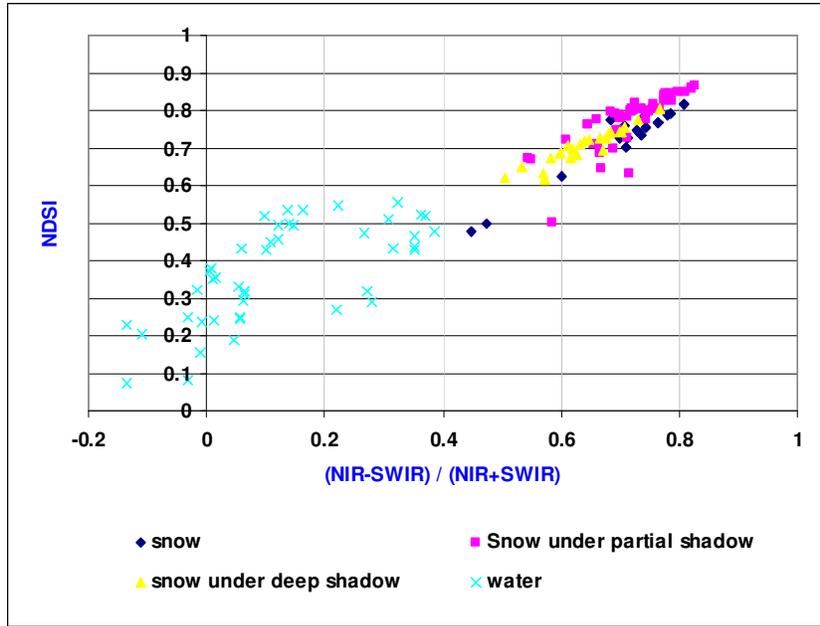


Fig. 5(a)

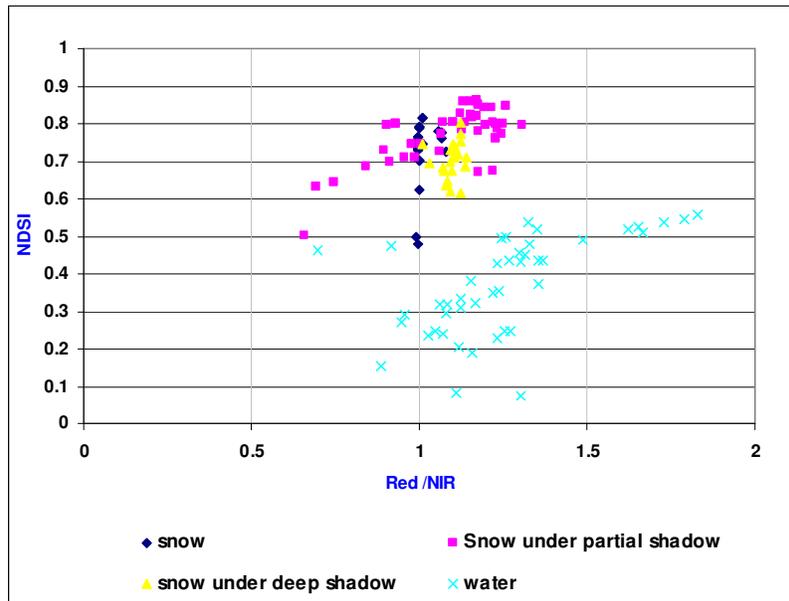
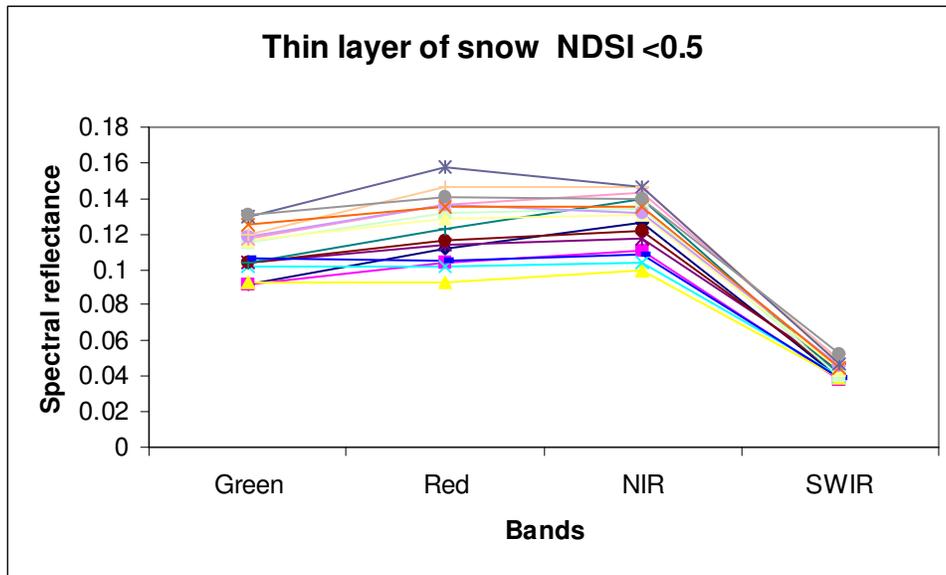


Fig. 5(b)

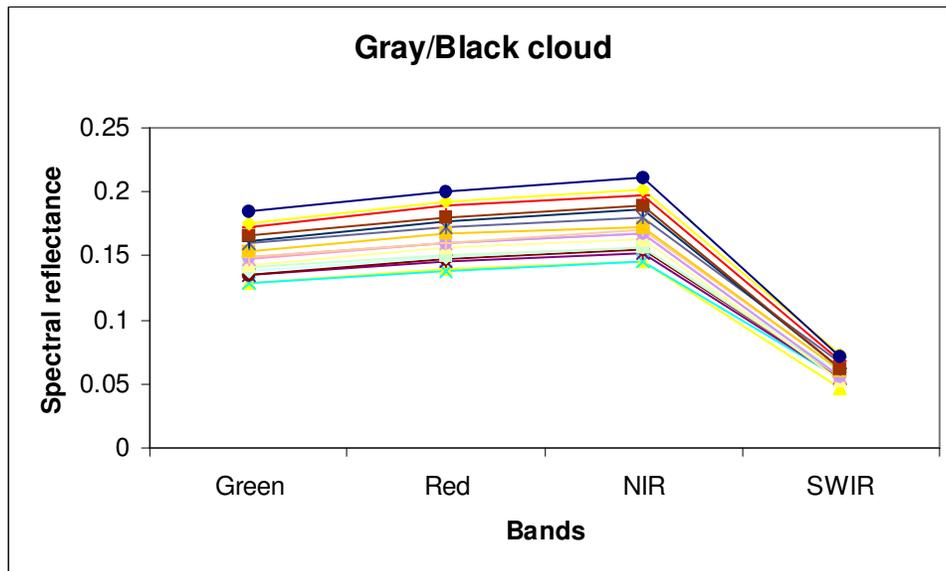
FIGURE 5 : Scatter plot NDSI for snow under mountain shadow and water as function of (a) normalized NIR –SWIR difference Index 9b) Normalized Red-NIR ratio

2.3.1. Elimination of Grey Cloud From Snow

It has been observed from scatter plot (Fig.3) that some of the pixels corresponding to black / grey clouds (water bearing clouds) has NDSI values >0.4 and <0.5 which may be due to the absorption of SWIR radiation by the water molecules present in this type of cloud. Hence, the spectral characteristics of the snow having $NDSI < 0.5$ and those of the gray (water bearing) cloud (Figure 6) has been studied to identify the discriminating characteristics. There exists small difference in the Green reflectance values, total pixel brightness and red-NIR ratio and these were used for the discrimination of the grey water bearing cloud and snow pixels having NDSI values between 0.4 and 0.5.



(a)



(b)

FIGURE 6 : The spectral characteristics of (a) snow having $NDSI < 0.5$; (b) gray (water bearing) cloud.

3. IMPLEMENTATION OF THE ALGORITHM

The automated snow cover extraction algorithm has been implemented in modular, hierarchical approach. The reflectance values for a number of carefully selected sample pixels were extracted from the image data acquired at different time periods and in different locations. Detailed statistical analysis has been carried out to arrive at suitable hierarchical logics and thresholds. At the top level of hierarchy, the candidate pixels for snow has been identified with $NDSI > 0.4$ and this has been implemented as shown below :

```
IF (NDSI>0.4) THEN {check snow logics}
ELSE {non snow pixel}
```

This logical condition eliminates the pixels corresponding to land cover classes such as barren soil, vegetation, urban areas and noise classes like bright cloud, but includes water and few dark cloud covered pixels.

Logics have been implemented for different NDSI ranges for discriminating the snow from those noise features interfere in that NDSI range. For example for $NDSI > 0.6$, noise feature-water has to be eliminated, while in the NDSI range below 0.5 that cloud pixels has to be eliminated in addition to the water. The exact feature parameter identification and the selection of the threshold value (or range) for each of the selected feature parameter has been determined based on statistical analysis of the spectral response. Typical analysis carried out for sub category-snow under shadow for few sample pixels used were given table 1.

A set of parameters viz., the normalized spectral indices, spectral ratios computed for arriving at the threshold values were shown in the table. Similar computations were made for each of sub categories of snow (under different levels of shadow) and the other features that introduce noise in the snow cover estimation such as cloud and water. The selection of the parameter and their threshold were determined by comparing mean, variance, minimum and maximum values of each of the sub category of snow and the mixing noise feature. The implementation of the hierarchical, multi threshold logics for the specific case (NDSI range : 0.5-0.6) has been implemented as given below :

```
IF ( ( NDSI >0.5) AND ( NIR-SWIR index>0.3) ) THEN
{
IF ( (NIR-SWIR index<0.57) AND ( R/NIR>1.15) AND ( Brightness < 0.25) )
THEN { Water}
Else IF (R/NIR < 1.0) THEN { snow }
}
Else IF( (NIR-SWIR index>0.40) AND ( Brightness > 0.375) )
THEN { snow }
}
```

Similar implementation has been carried out for each sub category of snow cover. In all about six such hierarchical set of logics were implemented in the new automated algorithm for snow cover estimation for the identification of snow and elimination of water and cloud pixels and is shown in the flowchart (Figure.7).

TABLE 1: Typical computation table for sub category – snow under shadow with few sample pixel data for the identification of feature parameter and their threshold values for use in hierarchical, multiple logics.

Green	Red	NIR	SWIR	NDSI	R/NIR	NIR-SWIR index	NDVI	sum of reflectance	G/R	G/NIR	R/SWIR	
0.203	0.187	0.175	0.026	0.77	1.07	0.74	-0.03	0.59	1.09	1.16	7.21	
0.203	0.187	0.175	0.026	0.77	1.07	0.74	-0.03	0.59	1.09	1.16	7.21	
0.201	0.188	0.272	0.045	0.63	0.69	0.72	0.18	0.71	1.07	0.74	4.17	
0.199	0.175	0.151	0.015	0.86	1.15	0.82	-0.07	0.54	1.14	1.31	11.72	
0.189	0.167	0.147	0.021	0.8	1.14	0.76	-0.06	0.52	1.13	1.28	8.17	
0.179	0.142	0.129	0.02	0.8	1.10	0.74	-0.05	0.47	1.27	1.39	7.26	
0.179	0.142	0.129	0.02	0.8	1.10	0.74	-0.05	0.47	1.27	1.39	7.26	
0.178	0.162	0.138	0.014	0.85	1.17	0.81	-0.08	0.49	1.10	1.29	11.24	
0.177	0.158	0.141	0.017	0.83	1.12	0.79	-0.06	0.49	1.11	1.25	9.41	
0.175	0.155	0.137	0.013	0.86	1.13	0.82	-0.06	0.48	1.13	1.28	11.56	
0.174	0.155	0.133	0.013	0.86	1.17	0.83	-0.08	0.47	1.12	1.31	12.23	
0.168	0.148	0.129	0.016	0.82	1.15	0.77	-0.07	0.46	1.13	1.30	9.07	
0.158	0.137	0.109	0.013	0.85	1.26	0.78	-0.11	0.42	1.15	1.45	10.43	
0.183	0.162	0.151	0.02	0.8	1.07	0.77	-0.03	0.52	1.13	1.21	8.15	
0.127	0.104	0.139	0.028	0.64	0.75	0.67	0.15	0.40	1.22	0.91	3.76	
0.108	0.095	0.105	0.019	0.7	0.91	0.69	0.05	0.33	1.13	1.03	4.95	
0.098	0.082	0.124	0.032	0.5	0.66	0.58	0.21	0.34	1.20	0.79	2.51	
0.079	0.064	0.052	0.015	0.67	1.22	0.55	-0.10	0.21	1.24	1.51	4.14	
0.078	0.062	0.053	0.015	0.67	1.18	0.55	-0.08	0.21	1.25	1.47	4.05	
0.183	0.162	0.151	0.02	0.76	1.06	0.73	-0.02	0.46	1.16	1.22	7.61	Mean
0.014	0.017	0.039	0.008	0.1	0.18	0.09	0.09	0.12	0.06	0.22	3.04	Std. Dev
0.203	0.188	0.272	0.045	0.86	1.26	0.83	0.21	0.12	1.27	1.51	12.23	Max
0.158	0.137	0.109	0.013	0.5	0.66	0.55	-0.11	0.12	1.07	0.74	2.51	Min

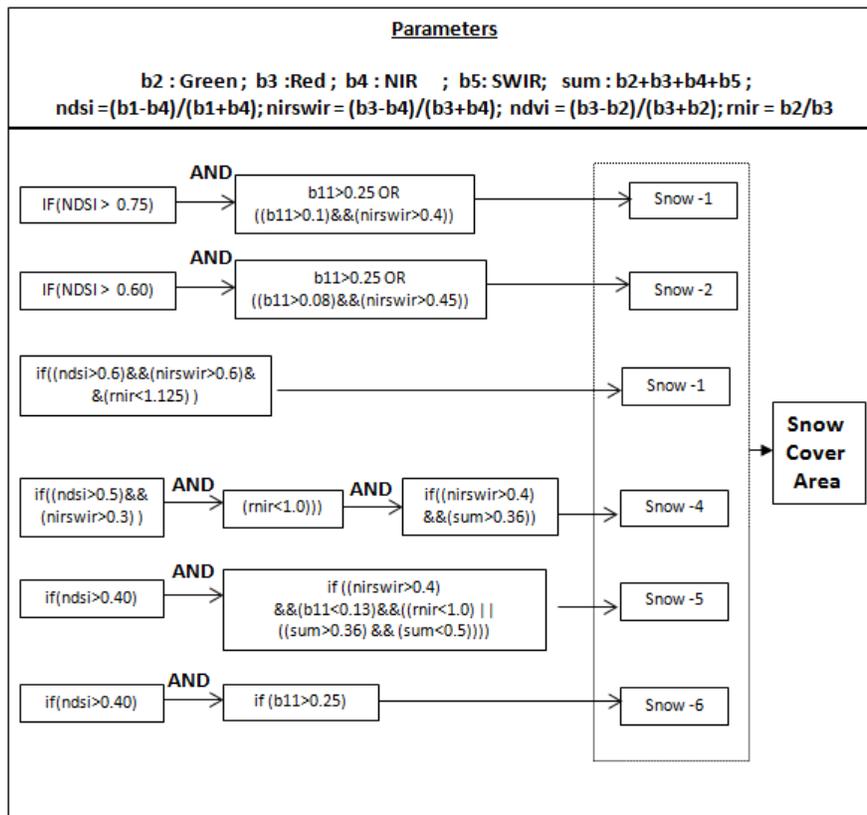


FIGURE 7 : The hierarchical logical equations used for automatic extraction of snow cover

4. VALIDATION OF THE RESULTS

The evaluation of the algorithm and validation of the results has been carried out by comparing the classification accuracy achieved with Hall’s algorithm and Kulkarni algorithm. The evaluation has been carried out using two approaches viz. sample pixel based analysis and the image based analysis. A number of sample pixels of the noise features such as cloud /water and different types of snow has been classified with all the three algorithms in the first approach and the results were shown in Figure. 8.

The new algorithm has been found to be more efficient in the elimination of noise features and is as good as Hall’s algorithm in the elimination of water pixels and has been able to almost eliminate the grey water bearing cloud pixels, which has not been addressed by either of the algorithms. The algorithm has also been found to be very effective in the detection of the snow pixels under shadow. These results has proved that the new snow algorithm is effective in eliminating the noise features in the automated snow cover estimation. A subset of AWiFS image which has noise features has been selected and the snow cover estimation has been

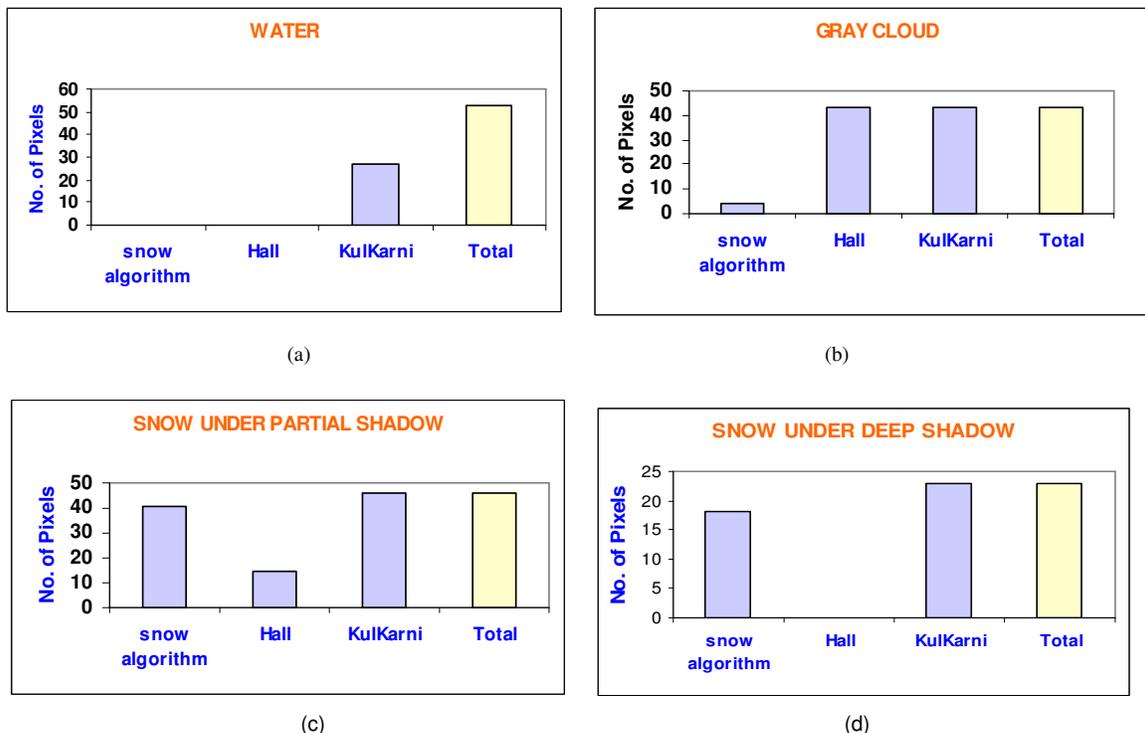


FIGURE 8: Elimination of (a) water (b) grey cloud ; inclusion of (c) snow under partial shadow (d) snow under deep shadow

To evaluate the performance of the algorithm in the exclusion of water bodies, a sub scene of AWiFS image (Fig 10a) has been used and the results obtained with the Hall's algorithm and that of the new algorithm are compared (figure 10b and 10c). The comparison of the results has shown that the new snow algorithm is as effective as that of the Hall's algorithm in the elimination of water pixels. Both the algorithm identifies the frozen water (ice) in a part of the Manasarover lake as snow cover pixels. This is due to the fact that the water in that part of the lake has frozen to form white ice. In this case, the Kulkarni's algorithm has not been considered for comparison as it has been expected to include all the water pixels having $NDSI > 0.4$. carried out using all the three algorithms and the results were evaluated for the improvements achieved with the new snow algorithm. Figure 9a. shows AWiFS image in snow covered mountain area of Himalayas, with a large number of snow cover pixels under mountain shadow. The levels of darkness / illuminations for these pixels vary widely. Most of the snow-covered areas under mountain shadow were eliminated in the snow cover classification results based on Hall's algorithm (figure 9b) due to the NIR reflectance threshold of $NIR > 0.11$. The snow cover estimation with new snow algorithm has included the snow pixels under shadow (Fig 9c) as expected. In case of Kulkarni's algorithm, all the snow cover pixels with $NDSI > 0.4$ will be detected and hence expected to provide results similar to that of new algorithm.

To evaluate the performance of the algorithm in the exclusion of water bodies, a sub scene of AWiFS image (Fig 10a) has been used and the results obtained with the Hall's algorithm and that of the new algorithm are compared (figure 10b and 10c). The comparison of the results has shown that the new snow algorithm is as effective as that of the Hall's algorithm in the elimination of water pixels. Both the algorithm identifies the frozen water (ice) in a part of the Manasarover lake as snow cover pixels. This is due to the fact that the water in that part of the lake has frozen to form white ice. In this case, the Kulkarni's algorithm has not been considered for comparison as it has been expected to include all the water pixels having $NDSI > 0.4$.

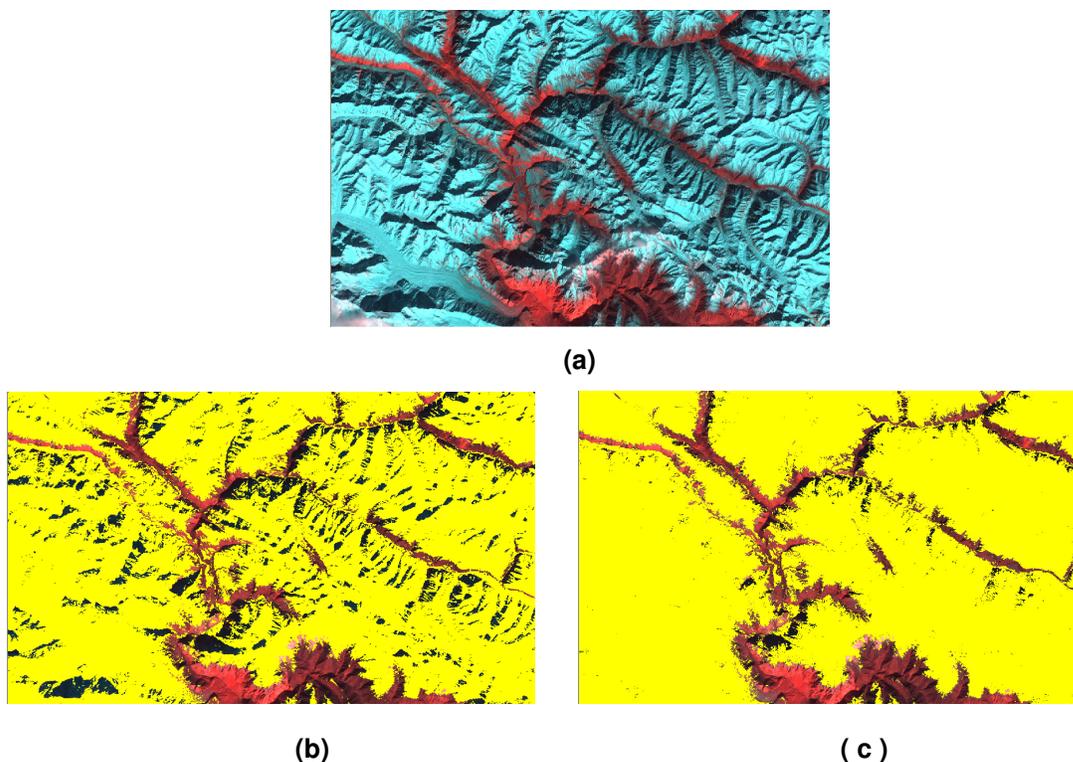


FIGURE 9: Extraction of snow cover under shadow conditions
(a) FCC image (B2, B3, B5) showing the snow cover (cyan); (b) FCC overlaid with snow cover derived with Hall's method (yellow); (c) FCC overlaid with snow cover derived with new snow algorithm (yellow)

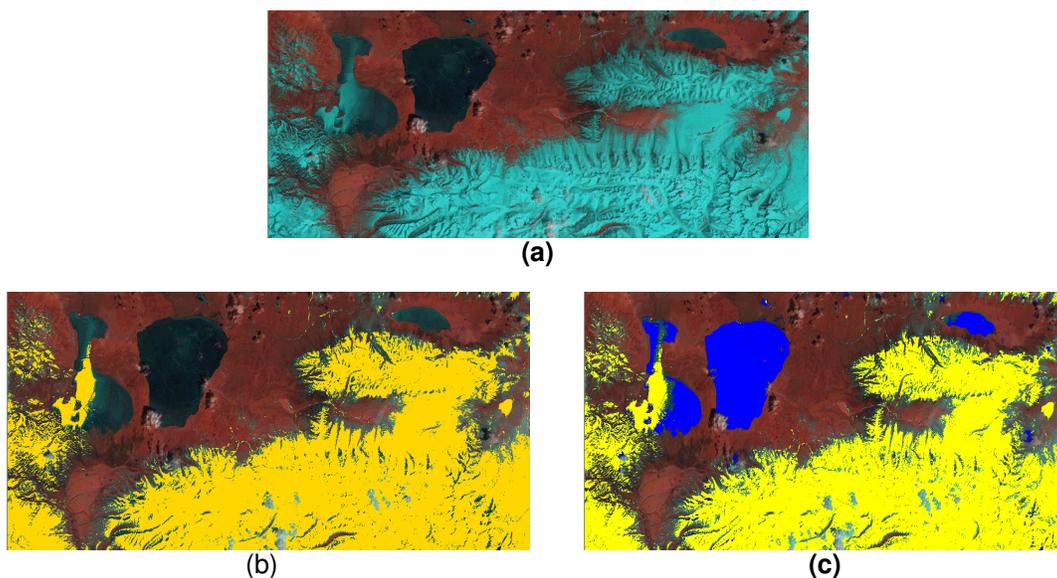
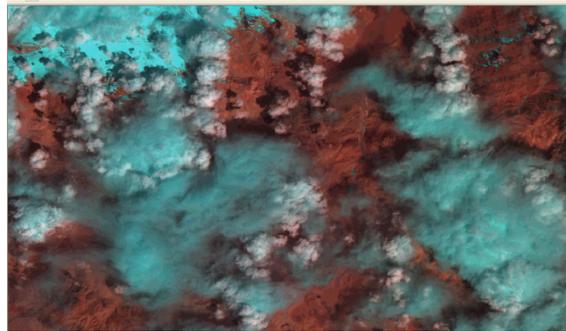
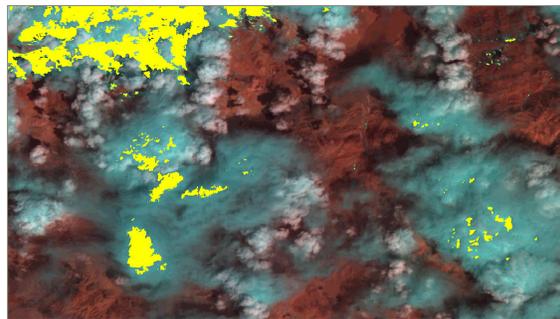


FIGURE 10 : Detection of snow in the presence of water bodies
(a) FCC sub-image(B2,B3,B5) showing the snow cover(cyan) ; (b) FCC overlaid with snow cover derived with Hall's method (yellow); (c) FCC overlaid with snow cover derived with new snow algorithm (yellow: snow ; Blue -Water)

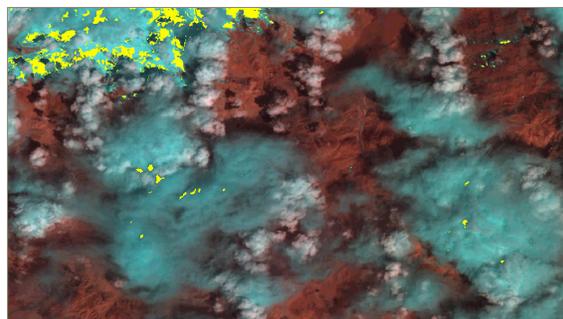
Figure 11a shows the AWiFS image with snow and gray (water bearing) clouds. The classification results obtained with the Kulkarni's algorithm and the new automated snow algorithm has been shown in figure 11b and 11c respectively. A substantial improvement in the reduction of cloud pixel noise has been observed in snow cover estimate by the new algorithm. The results of the Hall's threshold method is expected to be the same as that of Kulkarni's algorithm. This is due to the fact that the only additional condition for the Hall's method is the NIR reflectance has to be more than 0.11 and this condition has been satisfied for almost all the cloud pixels.



(a)



(b)



(c)

FIGURE 11: Elimination of clouds in snow cover extraction

(a) FCC sub-image(B2,B3,B5) showing the snow cover(cyan) ; (b) FCC overlaid with snow cover derived with Kulkarni's method (yellow); (c) FCC overlaid with snow cover derived with new snow algorithm (yellow)

6. RESULTS AND DISCUSSION

The snow cover estimation of the entire Himalayan mountain range covering India, Nepal, Pakistan and China has to be carried out for environmental studies related climate change. This implies that the algorithm for snow cover estimation has to handle larger image mosaic containing many AWiFS images. To validate the algorithm for its ability to handle image mosaics, it was tested with large image (42573 x 26881 pixels) covering Himalayan region of India from the state of Kashmir in the west to Arunachal Pradesh in the east (Figure 12). The algorithm has been found to be working satisfactorily. These results shows the robustness of the algorithm in various locations. Further it has been tested for the various seasons (October, Jan/Feb and April/May) covering the pre-snow fall period, winter and summer seasons.

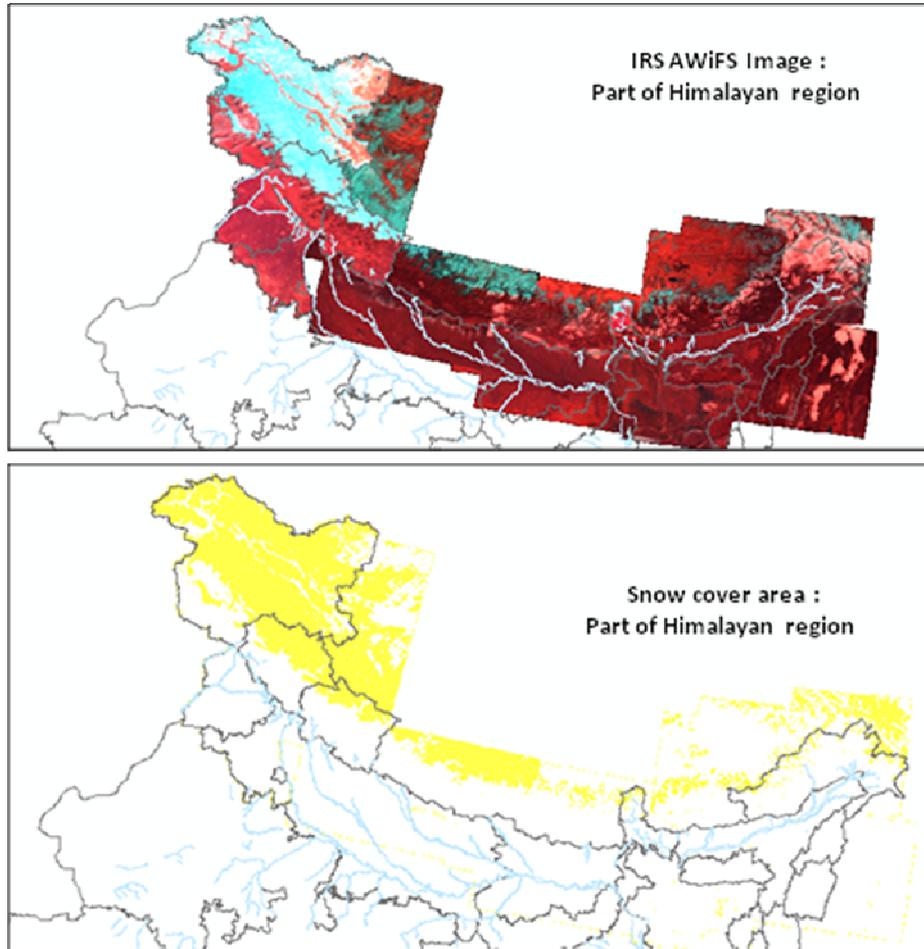
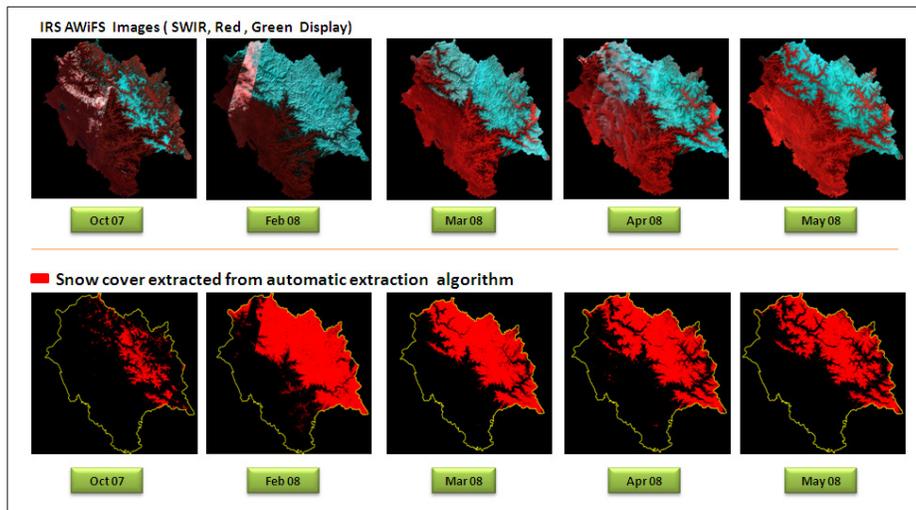
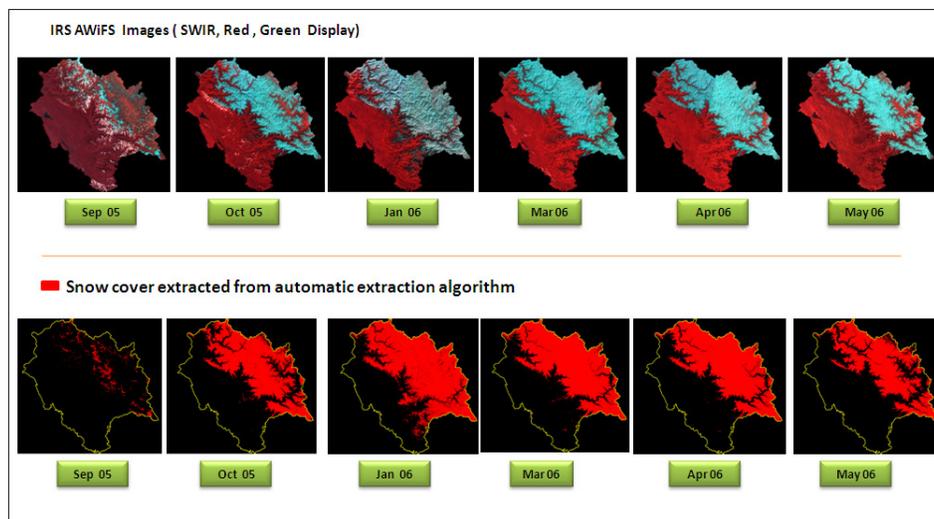


FIGURE 12: snow cover estimation results for AWiFS FCC mosaic (B2,B3 and B5) image scene of Himalayan region

The snow cover estimation for the years 2004-05 to 2007-08 has been evaluated over the Himalayan region of Indian terrain of the hilly mountain area in India and the results were ported into the NRSC web portal Bhoosampada [25]. The intra annual (temporal) change in the snow cover is an important parameter for the estimation of water flow form the snow melt in the rivers and the intra-annual studies are required in the case of environmental studies. Hence, an attempt has been made in the case of AWiFS imagery of Himachal Pradesh State, India for the years 2005-06 and 2007-08 as shown in Figure 13a. and the snow cover area estimates were shown in the Figure 13b. The maximum snow cover area occurs during the Jan-Feb (winter period) due to the snow fall that occurs in the Himachal Pradesh and it gradually reduces in the spring –summer period (March –May) to the extent of approximately 1/3 of the total snow covered area (Figure.14). It shows the minimum extent snow covered occurred in the September – October period before the onset of winter snow fall.



(a) 2007-08



(b) 2005-06

FIGURE 13. Temporal snow cover variation in Himachal Pradesh: (a) 2007-08; (b) 2005-06

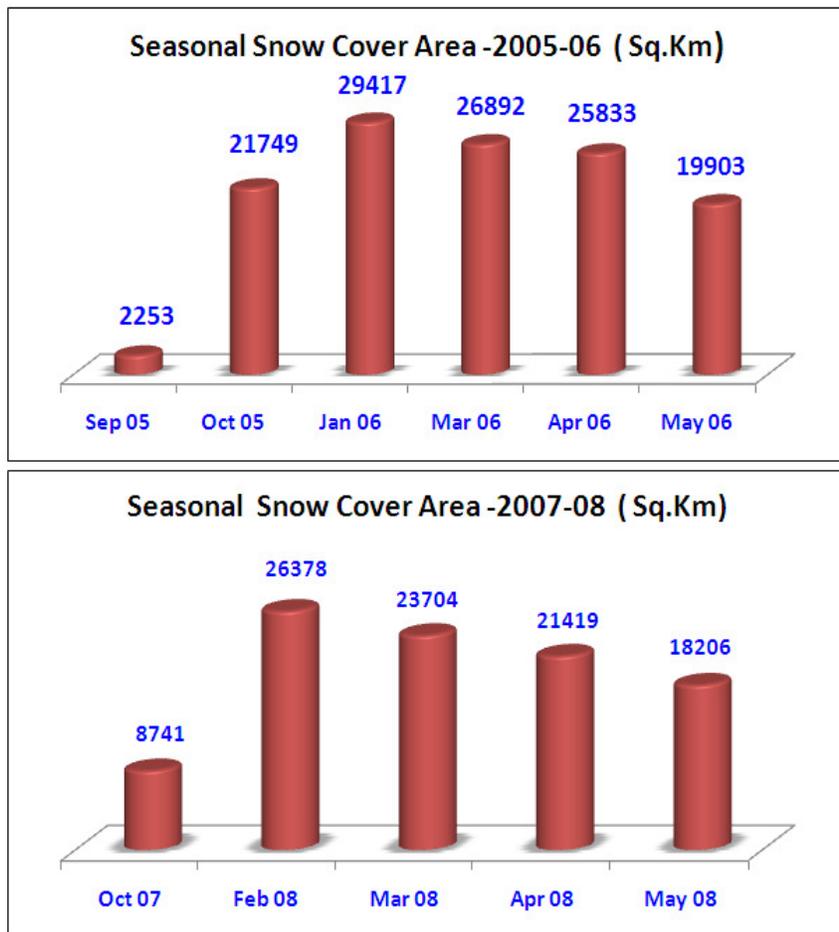
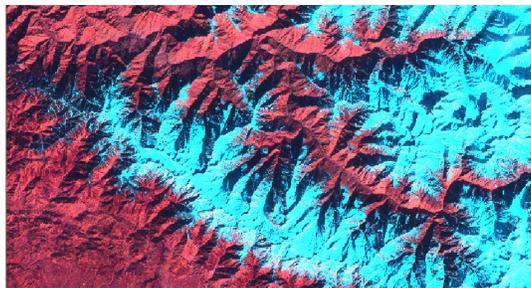


FIGURE 14. Seasonal snow cover area estimates

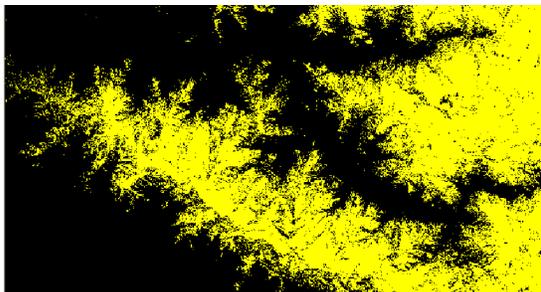
Since, automated snow cover extraction algorithm has been developed essentially based on the spectral contrast features of AWiFS, an attempt has been made to evaluate the same with the data obtained from other sensors having similar spectral bands such as Resourcesat –1 LISS III, LANDSAT ETM. However, it has to be noted that though, the spectral bands are similar, their radiometric resolutions are quite different (Table.2). The algorithm has been evaluated with the data obtained from IRS LISS III, Landsat TM data sets acquired on 02-02-2006, 02-06-2009 respectively were also shown in the Figure.15 and Figure 16. The results shows that the algorithm is suitable for these datasets as well. It has been observed a slight reduction in the snow cover area due to the lower radiometric resolution in ETM and LISS III. Hence, the required minor modifications for spectral thresholds were implemented to obtain the satisfactory result in the case of LISS III as shown in Figure. 15c. Similar exercise would improve the efficiency of the algorithm to enable the national / global application of the algorithm for the other sensors having similar spectral bands. The comparison of results obtained with LANDSAT ETM has shown the improvement achieved with the new algorithm compared to that of the Hall's method, particularly in the case of snow covered pixels in the mountain shadow region.

Satellite/ Sensor	Band width	Lmax	Lmin	DN max	Spatial Resolution	Radiometric Resolution
	μm	$\text{W/m}^2\text{-Sr-}$ micron	$\text{W/m}^2\text{-Sr-}$ micron		m	$(\text{W/m}^2\text{-Sr-micron})/$ Unit DN Value
Resourcesat -1 AWiFS						
B2	0.52 - 0.59	52.34	0	1024	56	0.051113281
B3	0.62 - 0.68	40.75	0	1024	56	0.039794922
B4	0.76 - 0.86	28.425	0	1024	56	0.027758789
B5	1.55 - 1.7	4.645	0	1024	56	0.004536133
Resourcesat -1 LISS III						
B2	0.52 - 0.59	12.064	0	128	23.5	0.094250000
B3	0.62 - 0.68	15.131	0	128	23.5	0.118210938
B4	0.76 - 0.86	15.757	0	128	23.5	0.123105630
B5	1.55 - 1.7	3.397	0	128	23.5	0.026539063
Landsat ETM						
B2	0.52-0.60	30.09	-0.64	256	30	0.120039063
B3	0.63-0.69	23.44	-0.05	256	30	0.096171875
B4	0.76-0.90	24.11	-0.51	256	30	0.096171875
B5	1.55-1.75	3.106	-0.1	256	30	0.012523438

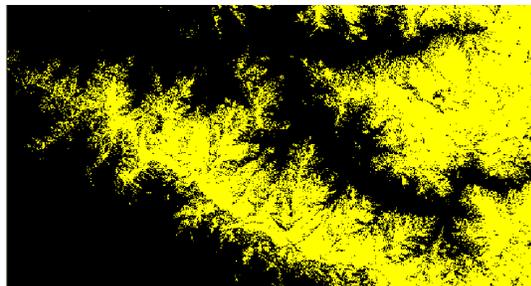
TABLE 2. Comparison of bandwidths in different satellite sensors in Visible & SWIR regions



(a)



(b)



(c)

FIGURE 15 : Algorithm results from IRS LISS III

(a) FCC sub-image(B2,B3,B5) showing the snow cover(cyan) ; (b) Snow cover derived from new automated algorithm for AWiFS ; (c) Snow cover derived from modified algorithm for LISS III

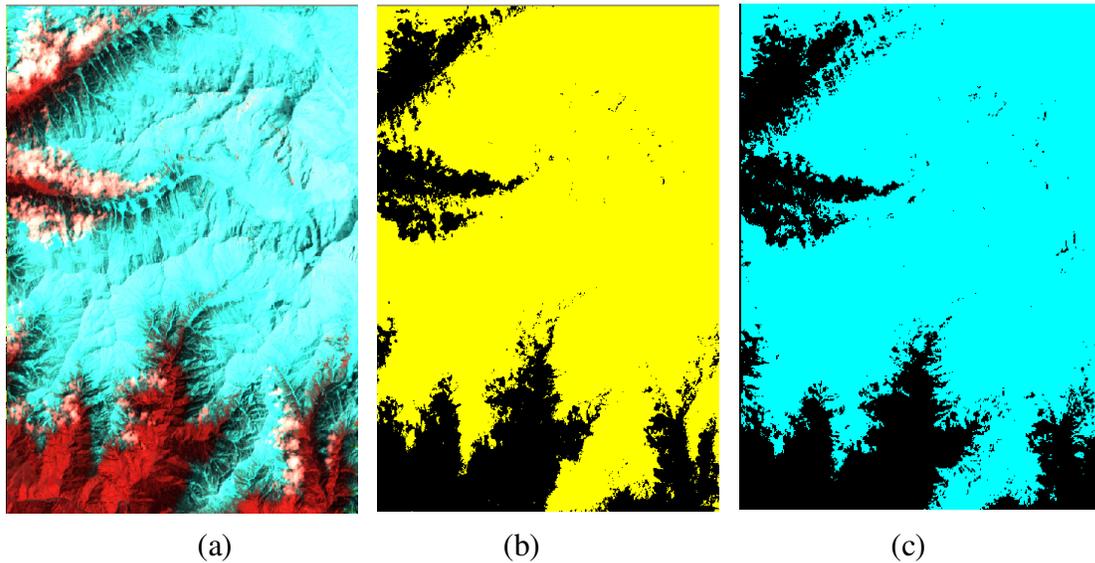


FIGURE 16. Algorithm results from Landsat ETM

(a) Landsat ETM sub-image(B2,B3,B5) showing the snow cover(cyan) ; (b) Snow cover derived from new automated algorithm ; (c) Snow cover derived from Hall's method

7. CONCLUSION

A new algorithm for the automated extraction of snow cover estimation has been developed and implemented for the Resourcesat-1 AWiFS sensor data. The algorithm has been developed based on the spectral signature of snow and other land cover noise in the spectral region of visible-NIR and SWIR spectral bands. The algorithm combines the best features of both Hall's algorithm and Kulkarni's algorithm overcoming their limitations. The algorithm has been evaluated extensively with large number of image data sets obtained over the Himalayas, bordering India in different seasons and is found to be working satisfactorily. The algorithm has been able to identify snow pixels in the presence of noise features like mountain shadow, water bodies and cloud (both white and water bearing gray cloud types). The algorithm has been found to be suitable for use with other sensor data such as IRS LISS III and Landsat ETM. The automatic snow cover estimation method can be applied for regular periodic monitoring of temporal snow cover change studies for snow melt runoff and estimation of water flow in the rivers originating in Himalaya and climate change. The seasonal snow cover over Himalayas in India over the 4 years (2004-05 to 2007-08) has been extracted from AWiFS data of IRS satellite and they are ported on to NRSC website BHOOSAMPADA (<http://applications.nrsc.gov.in:15001/>).

8. ACKNOWLEDGEMENTS

The authors are grateful to Director, NRSC for the motivation and kind support for the conduct of research work in the domain and guidance for implementation

9. REFERENCES

- [1] WWF Nepal Programme Report, 2005. "An overview of glaciers, glacier retreat and subsequent impact in Nepal, China and India". Available online at: http://www.panda.org/about_WWF/Climate_change.
- [2] A.V. Kulkarni, B.P. Rathore, S.K. Singh, and I.M. Bahuguna, 2011, "Understanding Himalayan cryosphere using remote sensing techniques" International Journal of Remote Sensing. Vol 32, pp 601-615.

- [3] Peru mountain glaciers 'receding rapidly' <http://www.enn.com/sci-tech/article/37794>
- [4] Jeff Dozzier.1989. "Spectral Signature of Alpine Snow Cover from Landsat Thematic Mapper". Remote sensing of Environment Vol 28 pp 9-22.
- [5] Jeff Dozzier. 1984. "Snow reflectance from Landsat-4 Thematic Mapper", IEEE Trans. Geosci. Remote Sensing. vol 22 pp 323-328.
- [6] D.K. Hall, G.A. Riggs and V.V. Satomanson, 1995, "Development of methods for mapping global snow cover using moderate resolution image spectroradiometer data", Remote sensing of Environment, vol. 54, pp 127-140.
- [7] D.K. Hall, J.L. Foster, D.L. Verbyla and A.G. Klein, 1998. "Assessment of snow-cover mapping accuracy in a variety of vegetation-cover densities in central Alaska", Remote sensing of Environment, 66, pp. 129-137.
- [8] R. Sidjak and R.D. Wheate,1999. "Glacier mapping of the Illecillewaet ice field, British Columbia, Canada, using Landsat TM and digital elevation data". International Journal of Remote Sensing, vol. 2, pp. 273-284.
- [9] D.K. Hall, Riggs, Salomanson, N.E. Digirolamno and K.J.Bayr, 2002. "MODIS snow cover products", Remote sensing of Environment, vol. 83, pp 181-194.
- [10] A.V. Kulkarni, J. Srinivasulu, S.S. Manjul and P. Mathur, 2002c. "Field-based spectral reflectance to develop NDSI method for snow coverage monitoring". Journal of the Indian Society of Remote Sensing, vol. 30, pp 73-80.
- [11] A.V. Kulkarni, S.K. Singh, P. Mathur and V.D. Mishra, 2006. "Algorithm to monitor snow cover using AWiFS data of RESOURCESAT-1 for Himalayan region", International Journal of Remote Sensing, vol. 27, pp 2449-2457.
- [12] X. Xiao, Z. Shen and X. Qin, 2001. "Assessing the potential of VEGETATION sensor data for mapping snow and ice cover: A normalized difference snow and ice index", International Journal of Remote Sensing, vol. 22, pp. 2479-2487.
- [13] X. Xiao, B. Moore, X. Qin, Z. Shen and Stephen Boles. 2002. "Large-scale observations of alpine snow and ice cover in Asia Using multi-temporal VEGETATION sensor data". International Journal of Remote Sensing, vol. 23, pp. 2213-2228.
- [14] Nando Foppa, Andreas Stoffel and Roland Meister, 2007, "Synergy of in situ and space borne observation of snow depth mapping in Swiss Alps", International Journal of Applied Earth Observation and Geoinformation, vol.9, pp 294-310.
- [15] Zhonghai jin, Thomas P. Charlock, Ping Yang, Yu Xie and Walter Miller.2008, " Snow Optical Properties for different particle shapes with application to snow grain size retrieval and MODIS.CERES radiance comparison over Antarctica", Remote sensing of Environment vol. 112, pp 3563-3581.
- [16] Miia Salminen, Jouni Pullianen Sari Metsamaki , Anna Kontu and Haane Suokanerva,2009. "The behavior of snow and snow-free surface reflectance in boreal forests: Implications to the performance of snow covered area monitoring", Remote Sensing of Environment vol.113, pp 907-918.

- [17] A.V. Kulkarni, S.S.Randhawa, B.P.Rathore, L.M.Bahuguna and R.K.Sood, 2002b. "A snow and glacier melt runoff model to estimate hydropower potential" . Journal of the Indian Society of Remote Sensing, vol. 30, pp 221-228.
- [18] V. Walter, 1998, "Automatic Classification of remote sensing Data for GIS Database Revision", IAPRS, vol. 32(4), pp 641-648.
- [19] A. Baraldi, V. Puzzolo, P. Blonda, L. Bruzzone and C.Tarantino, 2006, "Automatic spectral rule-based preliminary mapping of calibrated Landsat TM and ETM+ Images, IEEE trans. Geoscience and remote sensing, vol. 44(1) pp 2563-2586.
- [20] S. Subramaniam, A.V. Suresh Babu and P.S. Roy,2011, " Automated Water Spread Mapping Using ResourceSat-1 AWiFS Data for Water Bodies Information System", IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing, vol.4, pp 205-215.
- [21] D.Lu and Q. Weng. 2007, "A survey of image classification methods and techniques for improving classification performance", International journal of Remote Sensing, vol. 28, No. 5,pp 823-870.
- [22] M. Shoshany and Y. Cohen. 2007. "Inference verses evidence in reasoning remote sensing recognition with an information foraging", International journal of Remote Sensing, vol. 28, NO.11 pp 2613-2634
- [23] Amy E. Daniels, 2006, "Incorporating domain knowledge and spatial relationships into land cover classification: a rule based approach", International Journal of Remote Sensing, vol.27, NO.14, pp 2949-2975.
- [24] B.L. MARKHAM and I.J. BARKER.1986. "Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and at-satellite temperatures", EOSAT Technical Notes, vol. 1, pp. 3-8.
- [25] NRSC web portal : Bhoosampada (<http://applications.nrsc.gov.in:15001/>)

An Enhanced Image Warping Technique

Shakeel Doushant Imrith

*Department of Computer Science and Engineering,
University of Mauritius,
Reduit, Mauritius*

shakeel.imrith@umail.uom.ac.mu

Maleika Heenaye- Mamode Khan

*Department of Computer Science and Engineering,
University of Mauritius,
Reduit, Mauritius*

m.mamodekhan@uom.ac.mu

Abstract

Recently image warping is becoming a forefront subject and is attracting the attention of researchers. The motivation underpinning in exploring image warping is that it is producing wonderful effects on photographs and in film industries. Various warping algorithms are being devised to cater for the challenges posed by new image requirements. So far in literature, warping algorithms have been applied individually to produce pleasing effects. However, the amalgamation of several algorithms using appropriate proportions has been put aside. Furthermore, some existing warping algorithms lose some important features like contours in an image. In this paper, analysis of the mixture of morphing techniques has been applied on images to produce caricatures where the contours are cautiously preserved. The aesthetic effects of this newly devised amalgam algorithm is desirable to produce outstanding effects on face images.

Keywords: Image Morphing, Caricature, Blending.

1. INTRODUCTION

Nowadays, pictures are used everywhere, from the simple family picture from a digital camera to the 2 dimensional animations in the film industry. The pictures taken can experience optical distortions introduced by a camera or viewing perspective. There is an explosive need in correcting and manipulation these images to have a compelling and aesthetically pleasing effect on these pictures. Thus, image warping is a forefront subject which is gaining momentum [1, 2]. In fact, image warping is a process of geometrically transforming an image, based on pixel manipulation and color blending. It is in essence a transformation that changes the spatial configuration of an image [3, 4]. In the last decade, several warping techniques are being devised to quench the demand of film industries.

Many techniques have been devised in this field. Gooch et al. [5] used a 3 by 3 free form deformation in order to overstress the face image. Liang et al. [6] used the Partial Least Squares learning in order to predict the best pair if an image face was given. Last but not the least, there are several basic techniques like brightness, contrast, RGB (Red, Green, Blue), CMYK (Cyan, Magenta, Yellow, Black) and Hue, Saturation and Lightning which have been devised. These techniques have been amalgamated in order to form the new warping techniques which exist nowadays. For instance, contrast and brightness have been used in order to create techniques like noise or dithering [7]. Individually, these techniques hold little value but if they are combined in the right proportion, they may yield fantastic results. Xiong Bing et al. [8] demonstrated that orientation information, also termed as intensity level in the pixels can be used to map snakes effectively in a continuous way. Based on the same logic, if snakes can be continuous, it means that the contrast found on the contours forms a repeating pattern which draws the shapes of a face image. Moreover, as specified earlier, Gooch et al. have claimed that we recognize faces based on the difference in features present on them. Thus, they presented a caricature generator,

using a 3 by 3 deformation grid. Chiang et al. [9] exaggerated the most common features of a face, comparing it with the features of an average face. Chen et al. [10], on the other hand, did not over-stress upon the deformation of a face image in order to create a caricature. Instead, they learned art style from pairs of images and associated sketches drawn in a particular style. Their approach takes as input a face image and generates a well-shaped version of the image; that is, an elegant caricature. Freeman et al. [11] translated a sketch into different styles by translating the sketch through an example-based system. However, they focused on style transfer rather than generating a stylistic sketch from an image. Nguyen et al. [12] proposed the use of primitives in order to stylize the personality and structure of the portrait. They used the Golden Ratio, specified by Gooch et al. and further build on this to propose a novel decomposition of key exaggeration and suppression features, for constrained adaptation to cartoon templates. Wen et al. [13] have developed an image warping method based on level set. In this work, the image was warped locally and globally through level set dynamic evolution with a certain velocity field which controls the deformation.

Several warping techniques exist. So far, these traditional methods have been applied individually to generate effects on images. There are also many challenges like the disparity of contours in warped images that have not been attended to. Based on these facts, a new warping technique has been devised to produce enhanced caricature without fading the contour lines of the images. This new algorithm creates a black and white copy of a face image, based on the difference in contrast found in the contours of face images. The colors found on the face have different textures and hence different levels or degree of the same color; hence different contrast on different regions of the face. This new algorithm uses 5 of the basic warping techniques in order to deploy its effect on a face image. The results obtained after applying the fused devised algorithm provides more appealing results compared to individual algorithm. The rest of the paper is presented as follows. Section 2 gives an explanation on the basic warping techniques used. Section 3 gives the design of the new technique; that is, the proposed model. Section 4 gives the experimental results for this paper and lastly, section 5 gives the conclusion and the future works with respect to this paper.

2. WARPING TECHNIQUES

The warping techniques used to create the new technique are simple warping techniques, that is, modify the intensity and color space of an image. These techniques are based on the everyday life colors as well as the color models used in various part of our life, that is, in monitors, printers, for the human eye and for compression standards. To summarize, these techniques form part directly or indirectly of human psycho-visual aspect. As we know, human beings are more able to perceive colors in the form of Hue, Saturation and Brightness. However, other color models, such as RGB, CMYK, grayscale and contrast, have their importance. For instance, if contrast and CMYK are used in relevant proportions, it gives much weight to contrasted edges in any face image. In this section, the warping techniques applied on face images for the caricatures are explicitly explained.

2.1 RGB Preliminaries

The RGB color, used mostly in monitors, and human beings, being less prone to chrominance than luminance, are less able to recognize small changes in RGB color model. However, if this change reflects an impact on the intensity as well, the change, based on the psycho-visual ability of human, will be easily perceived. In this paper, we will use this technique in a proportion that it will yield a change in the contrast of the face image. In other words, it will help us to remove brighter colors and keep colors with more depth, so as to capture the edges of the face image. The main advantage in this is that it will enable us to modify the image without changing the aspect and features found in the image. Its disadvantage is that in small proportions, practically no changes are seen. The following figure shows the effects of RGB on images.

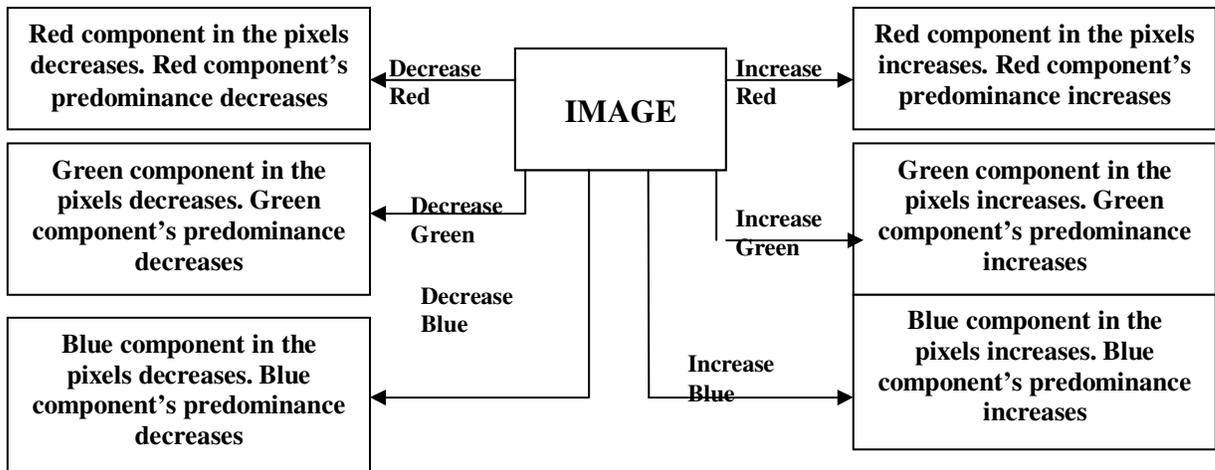


FIGURE 1: Effects of RGB on images



FIGURE 2: Example of Effects of RGB on images

2.2 CMYK Preliminaries

The CMYK color model is another technique used mainly for printing purposes where the black color needs to be dominant. This color model is of great help as it is a subtractive color model, as opposed to RGB. The black generated by this model will be used to generate the contours and edges of facial features, which are much contrasted in black color. This CMYK model will allow us to produce our final version of caricature by removing all the other colors and leaving the black color as the dominant color. In this way, the lighter parts of the face image will be removed. Thus, the main advantage of the CMYK model is its subtractive nature and its ability to make the contours of face images predominant. The results for the CMYK model are shown below.

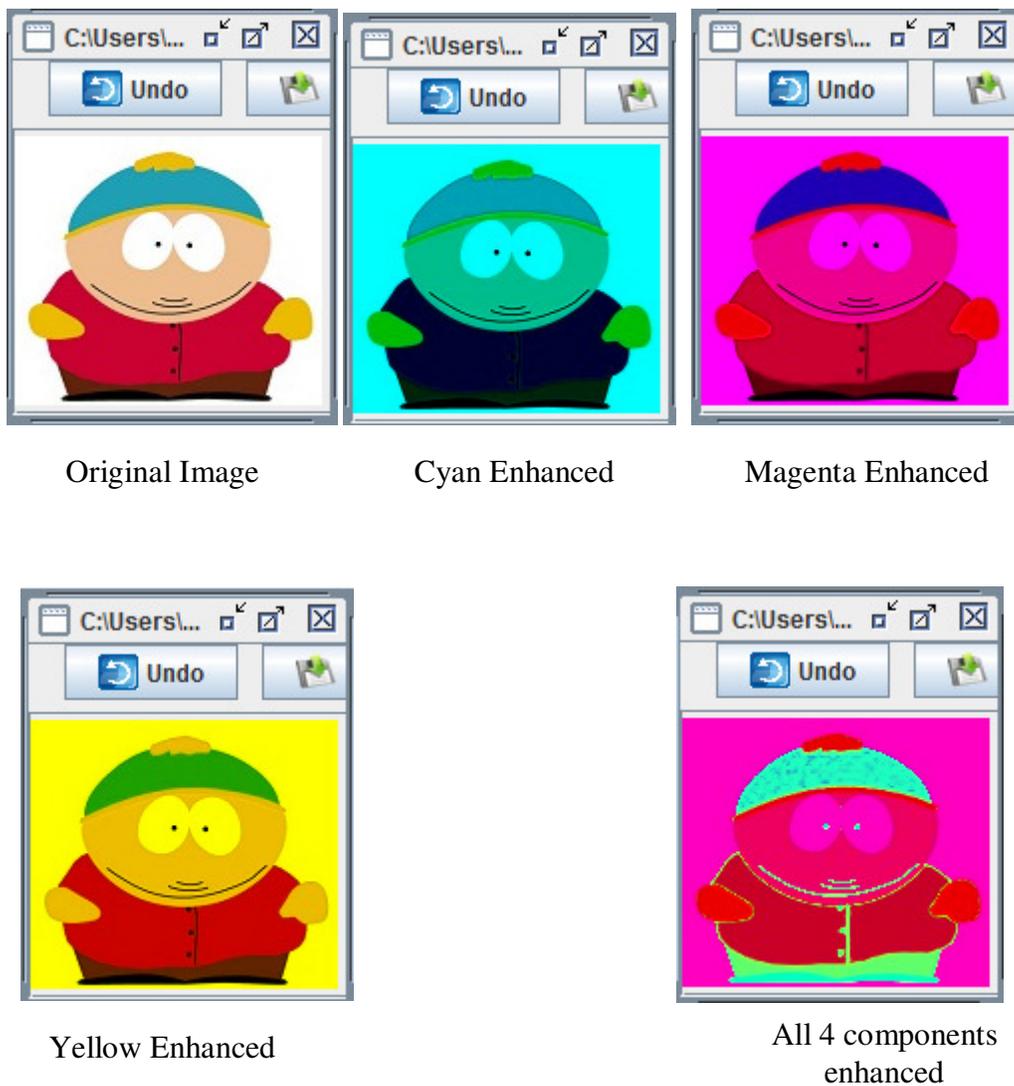


FIGURE 3: Example of Effects of CMYK on images

2.3 Grayscale Preliminaries

Another color model that is to be used is the grayscale color model. This technique helps in achieving uniformity of color in the pixels as well as allowing lighter pixels to be represented in light grey and contours to be represented in darker grey colors. The grayscale model allows us to remove different levels of colors pertaining to the same color so as to achieve regularity in terms of pixel intensity. The following figure shows the effect of grayscale on images.

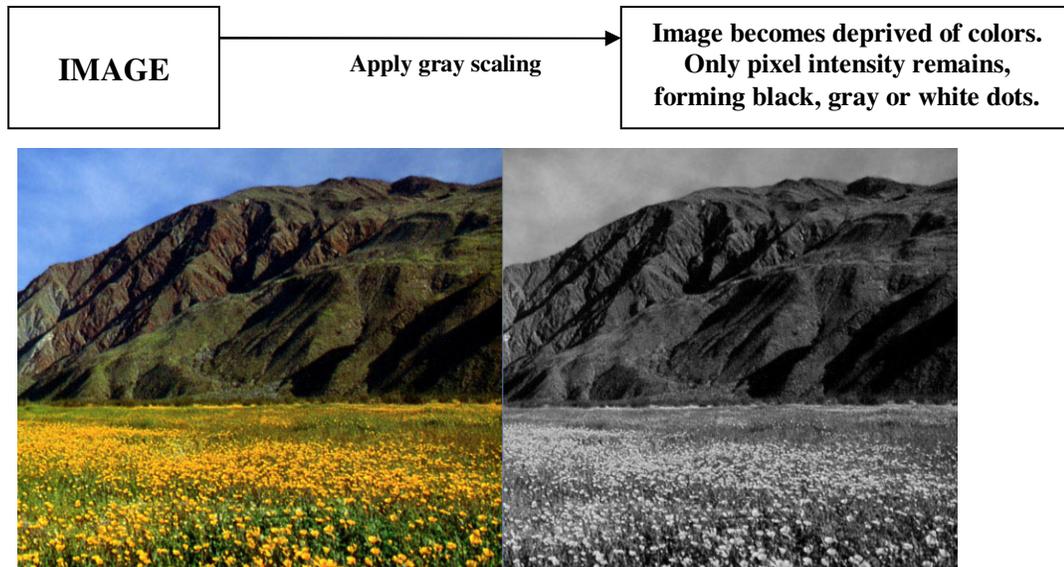


FIGURE 4: Example of Effects of grayscale preliminary on images

2.4 HSB Preliminaries

Human beings are more used to the hue, saturation and brightness color model. Hue is the attribute of a visual sensation according to which an area appears to be similar to one of the perceived colors such as red, green and blue. Saturation is the colorfulness of an area judged in proportion to its brightness. Brightness is the attribute of a visual sensation to which an area appears to emit more or less light. It is a fact that they have very little importance when used separately. However, when they are used together as a color model, they produce enhancing results. For instance, the hue will allow us to view the face image in different values of the RGB model. Saturation will help us to attenuate the colorfulness of the image, while the brightness will enable us to increase the pixel intensity of the pixels. The following figure shows the effect of applying HSB technique on images.

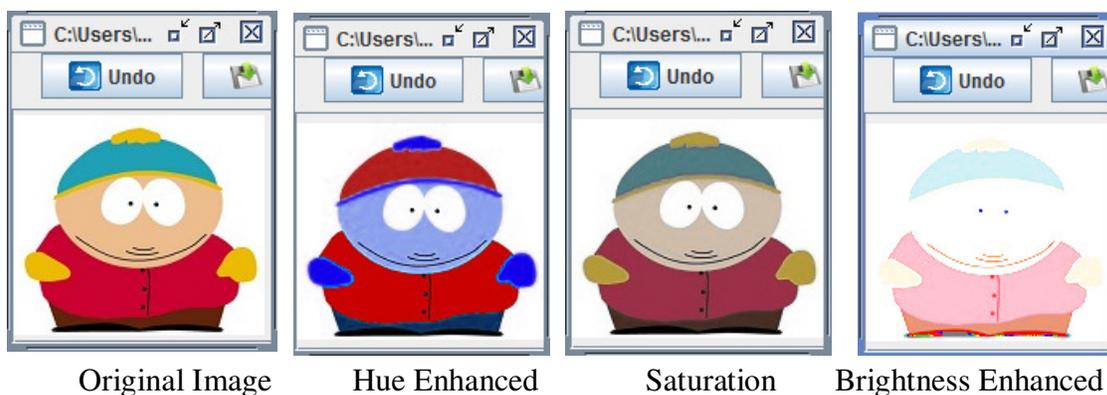


FIGURE 5: Example of Effects of HSB on images

2.5 Contrast Preliminaries

Another basic warping technique is the contrast model which allows us to inject some color disparity in contours or edges which are less defined due to photography errors. It enables features to be well shaped before the final touch of black and white drawing is applied to the face image.



FIGURE 6: Example of Effects of contrast preliminaries on images

2.6 Gamma Preliminaries

The gamma model is one which allows the uniform darkening of the pixels in an image. It helps in correcting a picture taken in very bright sunlight, by uniformly darkening pixels, irrespective of their respective intensities or the intensities of the surrounding pixels. On the other hand, if we decrease the gamma, it helps in making the image brighter, as shown below.



FIGURE 7: Example of Effects of Gamma preliminaries on images

These techniques discussed in this section are very basic and quite simple to implement. As illustrated in the above figures, they enhance contrast to some extent if they are used in the correct proportions. However, when they are applied individually, they give negligible enhancement to the face image like increasing the luminosity to some extent. Another issue is that the effect of warping techniques on images to produce caricatures is different. The following section gives an overview of different caricatures when applying warping techniques.

3. WARPING TECHNIQUES TO PRODUCE CARICATURES

As discussed in section 2.0, the effect of the basic warping techniques varies from one algorithm to the other. Special attention should be given to these techniques when building caricatures. In many cases, when caricatures are built, contours and edges are lost thus losing the appropriate presentation of the image. Caricature was used earlier to warp image and they were presented in the form of primitives, which are used to match features of any face image. The end result is just like applying the cartooning effect as shown in the figure below.



FIGURE 8: Image Caricature

Caricature can also be presented as the total deformation of face images. This was made possible by using the golden ratio and features and shapes were exaggerated. The end result was in the form of a puppet as shown in Figure 9.



FIGURE 9: Image Deformation

Last but not the least, caricature was termed in the form of polymorphing, that is, the use of different morphing algorithms on different face image in order to obtain the caricature as shown in figure 10.



FIGURE 10: Image Polymorphing

It is challenging to produce the caricatures as shown in figure 8, 9 and 10. The basic warping techniques that have been discussed in section 2 cannot produce such effect when applied individually. In fact, the essential use of the black color and ways of enhancing contrast was disregarded. While manipulating simple and basic pixel color and intensity manipulating

algorithm, enhancement in contrast can be made possible. Based on this enhancement, a black and white caricature can be obtained by removing all colors from any color space while keeping the black color in order to make the edges dominant. As mentioned earlier, there are scopes in developing more enhanced warping techniques to produce appealing effects on images. When the different techniques are used as an amalgam in different proportions, they yield about exquisite characteristics to face images. On this basis, a design is proposed and discussed in the next section.

4. IMPLEMENTATION OF PROPOSED WARPING TECHNIQUE

The enhanced algorithm devised consists of an amalgamation of 5 basic techniques namely: RGB, CMYK, grayscale, HSB and contrast preliminaries. However, the orders and proportion in which they are applied should be chosen cautiously. The following is the algorithm of the proposed warping technique:

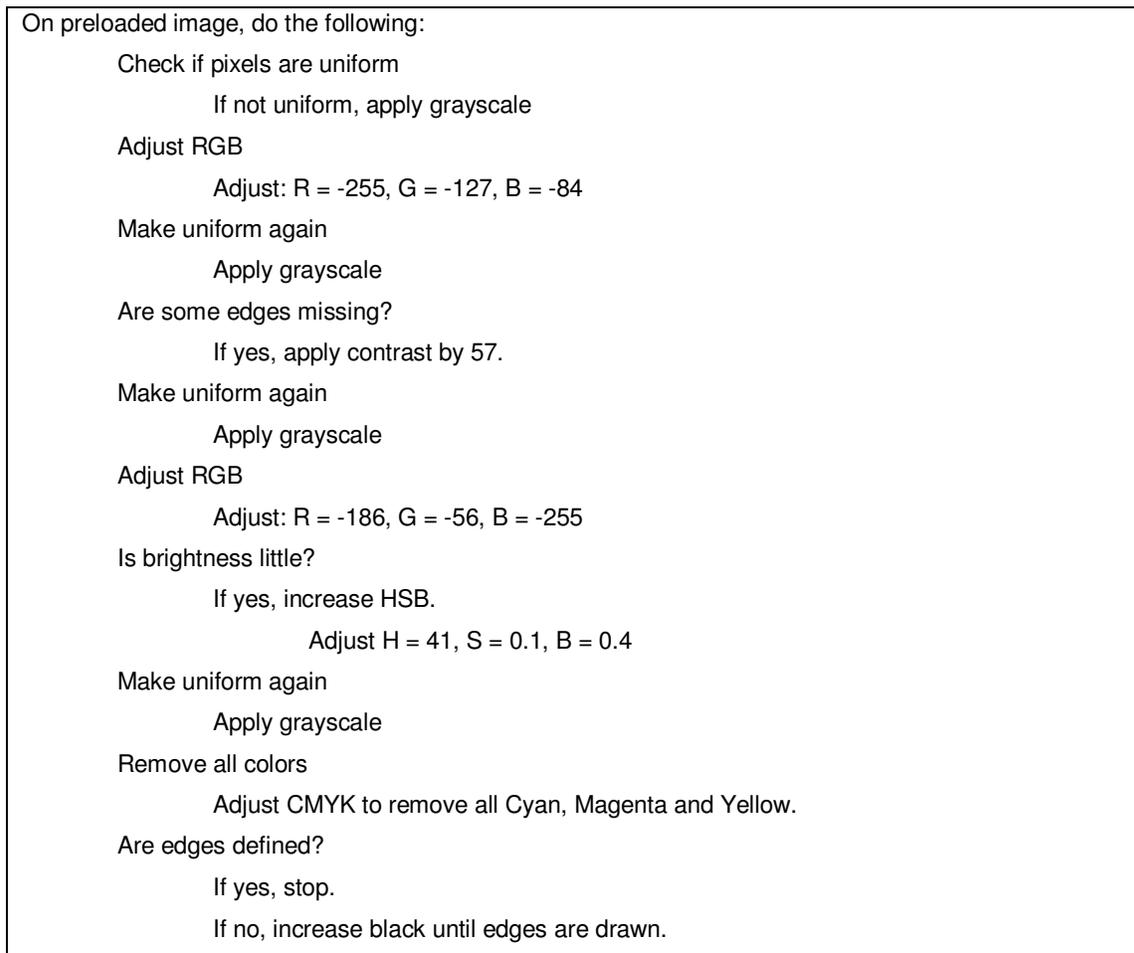


FIGURE 11: Newly devised algorithm

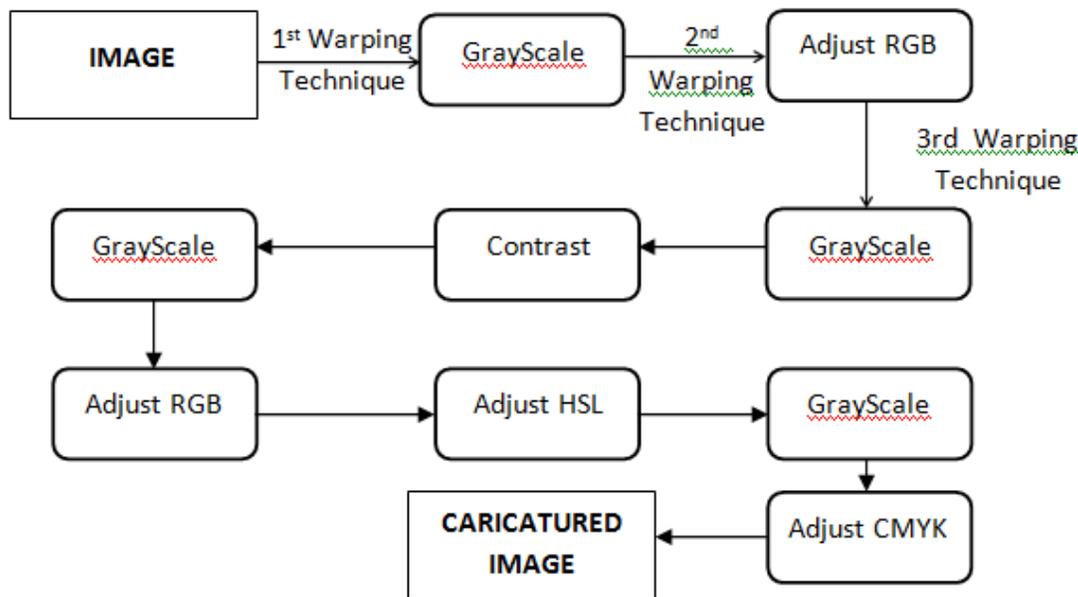


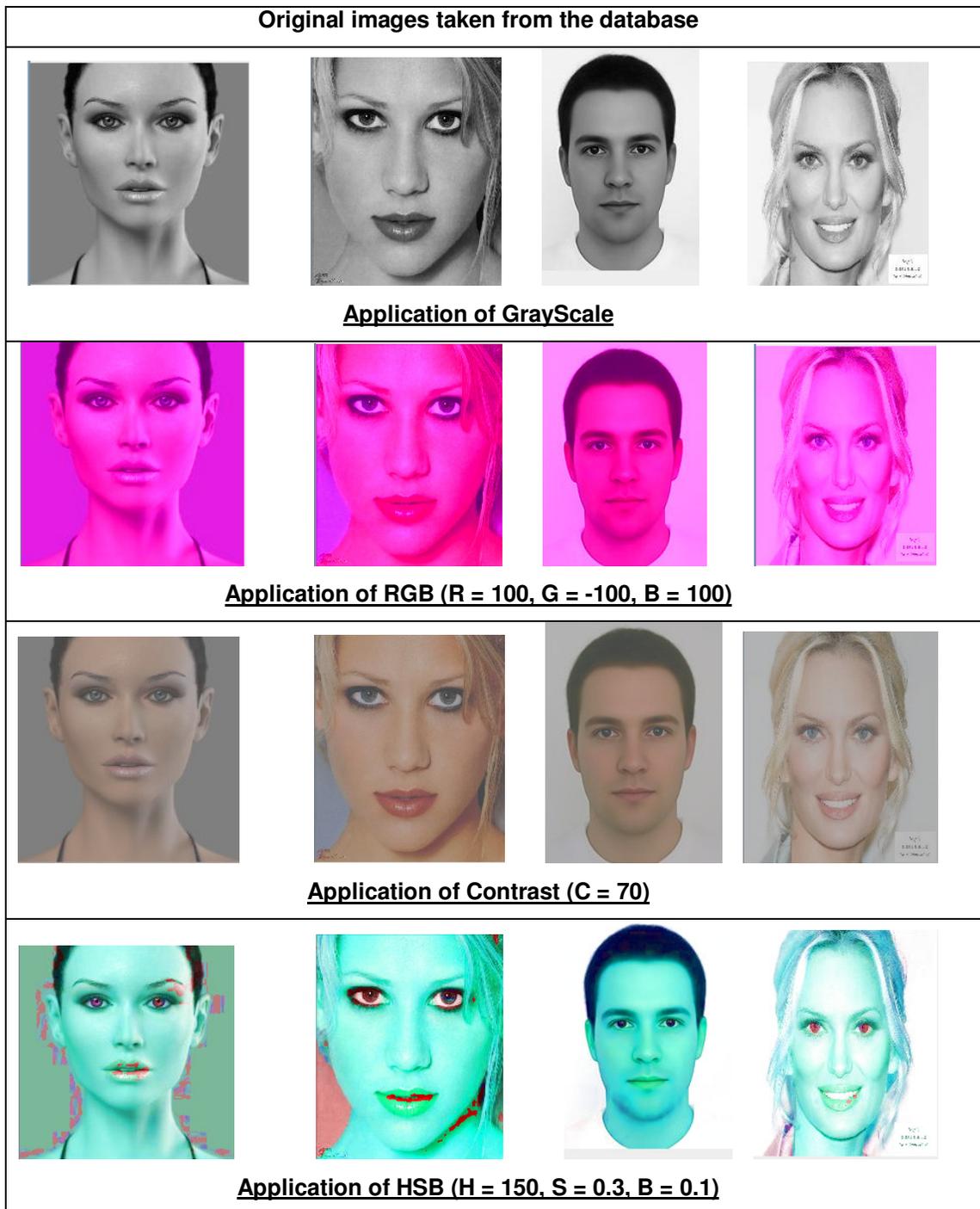
FIGURE 12: Newly devised design

First of all, there is a need to bring about uniformity in pixel levels and remove colors. The second step is used to darken the image so that lighter intensity pixels are removed in subsequent steps. The third step helps to bring about some uniformity again to the face image. The fourth step will clarify the image a bit and help in defining the contours well. Contours which were already well defined are amplified, while contours, which lack contrast, are given a boost. The fifth step may not yield any change to the current image but its aftermath cannot be disregarded. The sixth step will darken the image beyond recognition but the features are well distinct from the rest of the image. The seventh step deals with the human psycho-visual ability and will help in modifying the RGB space, decrease colorfulness of the image and brighten it so as to see the contours well defined. Uniformity is applied again in the eighth step to unify the colors brought about by the seventh step. Finally the last step will remove all colors from the pixels and inject some black color only to the image. This will make pixels of light intensity to disappear, which results in sharper and nicer contours of the face image.

A database of different face images were used to experiment with the newly devised algorithm and some of the results are shown in the following section.

5. EXPERIMENTAL RESULTS

The first table shows the result of the individual warping algorithms, that is, Grayscale, RGB, Contrast, HSB and CMYK applied on 4 different face images taken from the face databases.



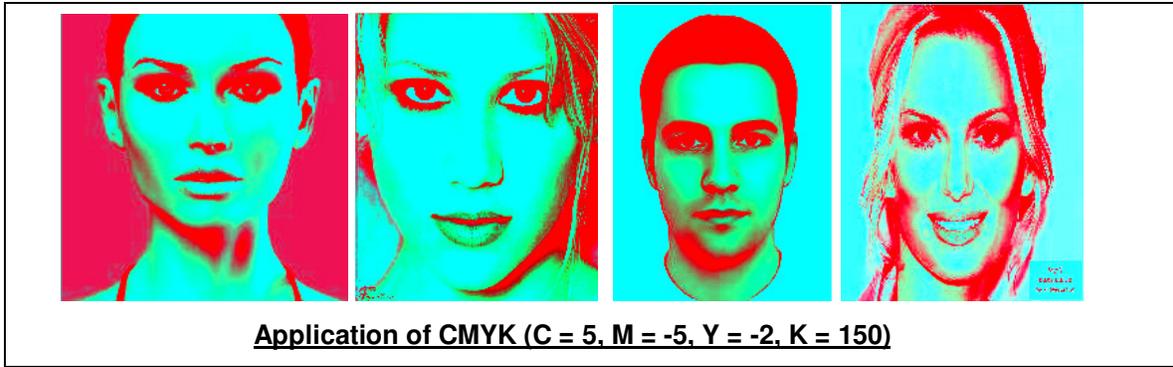
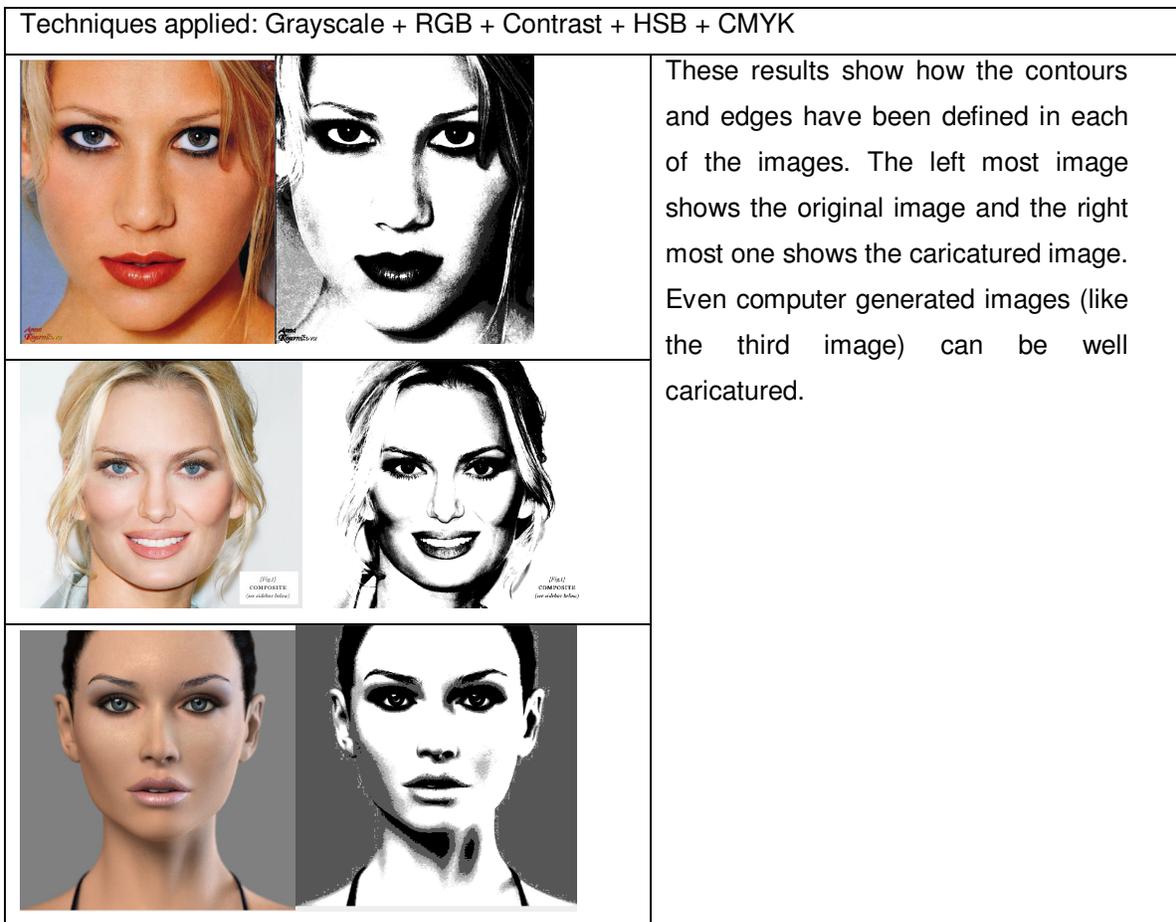


FIGURE 13: Application of gray scale, RGB, Contrast, HSB and CMYK individually

Figure 12 shows the effect of warping algorithms applied individually on face images. The effects are different when applying the newly devised algorithms presented in this paper. All the edges and the contours are preserved making the algorithm an ideal for caricatures. Figure 13 shows the face images after applying the algorithm presented in this paper.



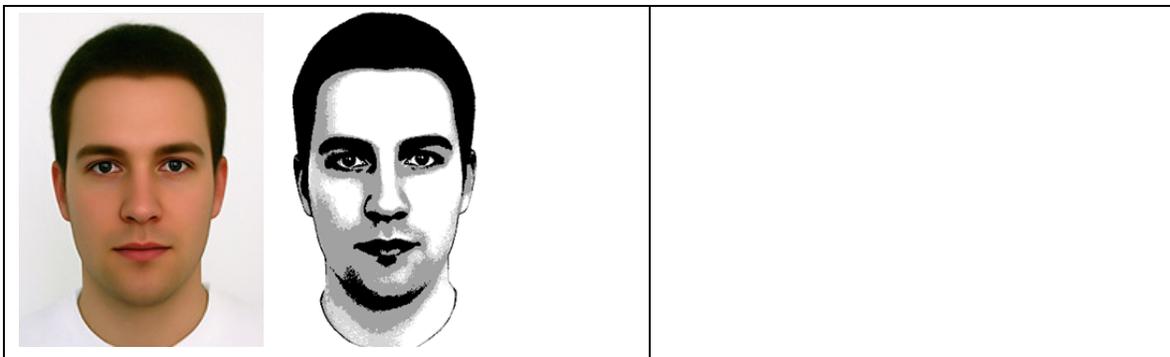
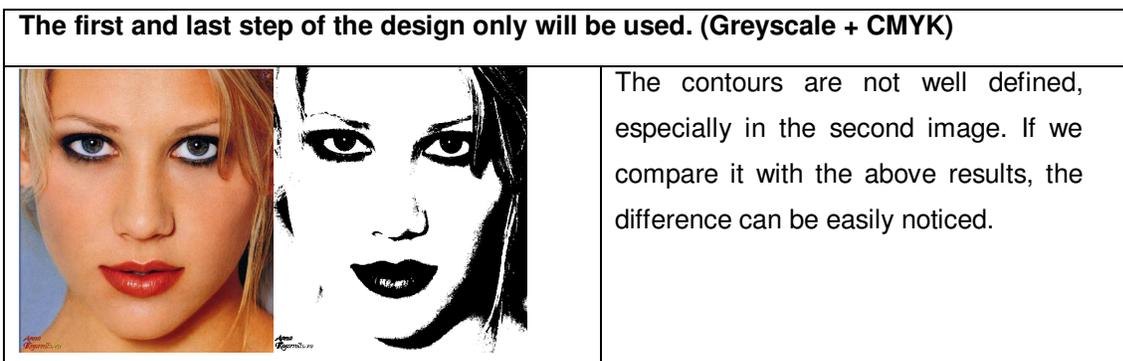


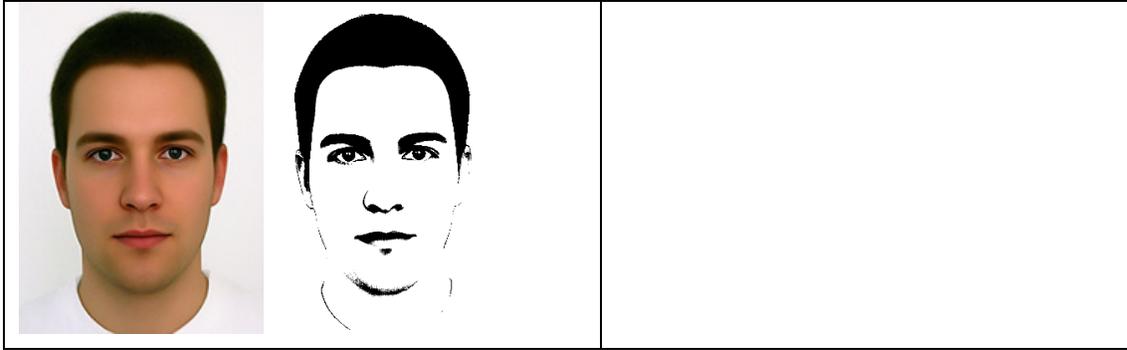
FIGURE 14: Application of newly devised algorithm

In the above example, the caricature technique in section 4 was applied in the following proportions:

- Grayscale
- RGB applied: R = -255, G = -127, B = -84
- Grayscale
- Contrast applied: Contrast = 57
- Grayscale
- RGB applied: R = -186, G = -56, B = -255
- HSL applied: H = 41, S = 0.1, L = 0.4
- Grayscale
- CMYK applied: C = -25, M = -25, Y = -25, K = Adjust at will until nice effect is observed.

The devised application was also tested with some missing warping algorithms among the five algorithms to analyze the effect that they have on the face images. It can be noticed that contours are not well defined leading to the loss of some features representing the images or the colors are not properly defined.





The first, second and last step of the design only will be used. (Greyscale + RGB + CMYK)



Contours appear but there seems to be some amalgamation of colors. Moreover, a portion of blue color appears.



The first, second, seventh and last step of the design only will be used. (Greyscale + RGB + HSB + CMYK)



Contours appear but there seems to be some amalgamation of colors. Moreover, a portion of blue color appears.

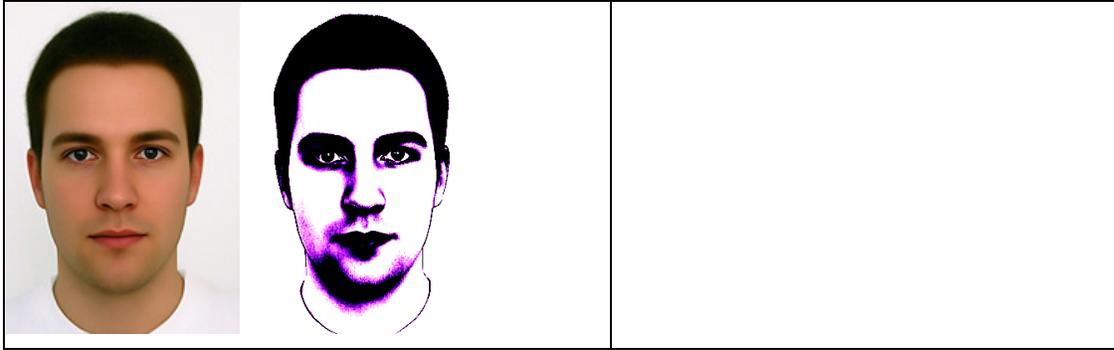


FIGURE 15: Results when some warping algorithms are omitted

6. CONCLUSION

In this paper, an enhanced warping algorithm is devised based on the basic pixel and image warping techniques. The right proportion of any basic technique will yield a very good finish. At first, the use of basic warping and pixel manipulation techniques seemed trivial. However, as they were mixed in several proportions, an exquisite result appeared. Caricature is used in many spheres of life. To build well-defined caricatures appropriate warping techniques should be used. Nowadays, we can make a portrait of ourselves in just a click of a mouse. When we knew how edges and contrast were related, using the right techniques in the right proportions allowed us to exploit this aspect of face image, which yielded to caricature.

7. REFERENCES

- [1] Wolberg, G., Recent advances in image morphing, International presentation of Computer Graphics, June 1996, Pohang, Korea.
- [2] Mikkel B Stegmann, *Image Warping* Informatics and Mathematical Modelling, Technical University of Denmark, DTU, October 2001
- [3] Stoffels S., "Full mesh warping techniques", Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-00), Verona, Italy, December 7-9, 2000.
- [4] Glasbey C.A., Mardia K.V., "A review of image warping methods", Journal of Applied Statistics, 25, 155-171, 1998.
- [5] Gooch B., Reinhard E., Gooch A., "Human facial illustrations: Creation and psychophysical evaluation", ACM Trans. Graph., 23(1), pp. 27-44, 2004.
- [6] Liang L., Chen H., Xu Y-Q, Shum H-Y, "Example-based Caricature Generation with Exaggeration", IEEE Proceedings of the 10th Pacific Conference on Computer Graphics and Applications, 2002.
- [7] J. Shyh-Kang, "Audio Signal Processing-Quantization", Department of Electrical Engineering/ Graduate Institute of Communication Engineering.
- [8] X. Bing, Y. Wei and C. Charoensak, "Face contour tracking in video using active contour model", School of Electrical and Electronic Engineering, Nanyang Technological University, IEEE 2004.
- [9] Chiang P-Y, Liao W-H, Li T-Y, "Automatic Caricature Generation by Analyzing Facial Features", 2004 Asian Conference on Computer Vision, Korea, Jan 27-30, 2004.

- [10] Chen H., Xu Y., Shum H., Zhu S., and Zheng N., "Example based facial sketch generation with non-parametric sampling", ICCV01, pages II: 433–438, 2001.
- [11] Freeman W. T., Tenenbaum J. B., and Pasztor E., "An example-based approach to style translation for line drawings", Technical Report 11, MERL Technical Report, Cambridge, MA, Feb 1999.
- [12] Nguyen K. H. L, Yong P. W., Golam A., "Shape Stylized Face Caricatures", School of Computing, National University of Singapore.
- [13] Wen P., Wu, X., Huang W., Ren R., "An image warping method based on level set", 2nd International Congress on Image and Signal Processing, Pg. 1569-1573, 2010

Comprehensive Performance Comparison of Cosine, Walsh, Haar, Kekre, Sine, Slant and Hartley Transforms for CBIR with Fractional Coefficients of Transformed Image

Dr. H. B. Kekre

Senior Professor, Computer Engineering,
MPSTME, SVKM'S NMIMS (Deemed to be University),
Mumbai, 400056, India

hbkekre@yahoo.com

Dr. Sudeep D. Thepade

Associate Professor, Computer Engineering,
MPSTME, SVKM'S NMIMS (Deemed to be University),
Mumbai, 400056, India

sudeepthepade@gmail.com

Akshay Maloo

Systems Engineer, Infosys Limited,
Pune, India

akshaymaloo@gmail.com

Abstract

The desire of better and faster retrieval techniques has always fuelled to the research in content based image retrieval (CBIR). The extended comparison of innovative content based image retrieval (CBIR) techniques based on feature vectors as fractional coefficients of transformed images using various orthogonal transforms is presented in the paper. Here the fairly large numbers of popular transforms are considered along with newly introduced transform. The used transforms are Discrete Cosine, Walsh, Haar, Kekre, Discrete Sine, Slant and Discrete Hartley transforms. The benefit of energy compaction of transforms in higher coefficients is taken to reduce the feature vector size per image by taking fractional coefficients of transformed image. Smaller feature vector size results in less time for comparison of feature vectors resulting in faster retrieval of images. The feature vectors are extracted in fourteen different ways from the transformed image, with the first being all the coefficients of transformed image considered and then fourteen reduced coefficients sets are considered as feature vectors (as 50%, 25%, 12.5%, 6.25%, 3.125%, 1.5625%, 0.7813%, 0.39%, 0.195%, 0.097%, 0.048%, 0.024%, 0.012% and 0.06% of complete transformed image coefficients). To extract Gray and RGB feature sets the seven image transforms are applied on gray image equivalents and the color components of images. Then these fourteen reduced coefficients sets for gray as well as RGB feature vectors are used instead of using all coefficients of transformed images as feature vector for image retrieval, resulting into better performance and lower computations. The Wang image database of 1000 images spread across 11 categories is used to test the performance of proposed CBIR techniques. 55 queries (5 per category) are fired on the database to find net average precision and recall values for all feature sets per transform for each proposed CBIR technique. The results have shown performance improvement (higher precision and recall values) with fractional coefficients compared to complete transform of image at reduced computations resulting in faster retrieval. Finally Kekre transform surpasses all other discussed transforms in performance with highest precision and recall values for fractional coefficients (6.25% and 3.125% of all coefficients) and computation are lowered by 94.08% as compared to Cosine or Sine or Hartley transforms.

Keywords: CBIR, Image Transform, DCT, Walsh, Haar, Kekre, DST, Slant, Hartley, Fractional Coefficients.

1. INTRODUCTION

The computer systems have been posed with large number of challenges to store/transmit and index/manage large numbers of images effectively, which are being generated from a

variety of sources. Storage and transmission is taken care by image compression with significant advancements been made [1,4,5,43]. One of the promising and important research area for researchers from a wide range of disciplines like computer vision, image processing and database areas is image indexing and retrieval [2,6,7,10,11]. The desire of better and faster image retrieval techniques is till enticing to the researchers working in some of important applications for CBIR technology like museums, archaeology [3], art galleries [12,14], weather forecast [5,22], architecture design [8,13], geographic information systems [5], criminal investigations [24,25], medical imaging [5,18], trademark databases [21,23], image search on the Internet [9,19,20].

1.1 Content Based Image Retrieval (CBIR)

The interest in CBIR is growing because of the limitations inherent in metadata-based systems, as well as the large range of possible applications for efficient image retrieval. In literature Kato et. al. [4] used the term content based image retrieval (CBIR) for the very first time, to describe his experiments on automatic retrieval of images from a database by colour and shape feature. Textual information about images can be easily searched using existing technology, but requires humans to personally describe every image in the database. This is impractical for very large databases, or for images that are generated automatically, e.g. from surveillance cameras. It is also possible to miss images that use different synonyms in their descriptions. Systems based on categorizing images in semantic classes like "cat" as a subclass of "animal" avoid this problem but still face the same scaling issues [9,19].

Mainly two major tasks are performed by CBIR system [16,17]. Feature extraction (FE) is the first one, where a set of features, called feature vector, is generated to accurately represent the content of each image in the database. Similarity measurement (SM), is the second one where a distance between the query image and each image in the database using their feature vectors is used to retrieve the "closest" images [16,17,26]. For CBIR feature extraction the two main approaches are feature extraction in spatial domain [5] and feature extraction in transform domain [1]. The feature extraction in spatial domain includes the CBIR techniques based on histograms [5], BTC [2,16,23], VQ [21,25,26]. As the transform domain methods are widely used in image compression, as they give high energy compaction in transformed image[17,24], hence images in transformed domain are used for feature extraction in CBIR [1]. The energy compaction in few elements, so large number of the coefficients of transformed image can be neglected to reduce the size of feature vector [1].

Using individual image transforms various CBIR techniques have been proposed, but so far comparison of transforms is not being studied thoroughly [41,42]. In the paper the performance of fairly large number of popular transforms for CBIR are compared along with newly introduced Kekre transform. Getting the improvement in the performance of image retrieval technique even with reduced size feature vector using fractional coefficients of transformed image is the theme of the work presented here. Many current CBIR systems use average Euclidean distance [1,2,3,8-14,23] on the extracted feature set as a similarity measure. The direct Average Euclidian Distance (AED) between image P and query image Q can be given as equation 1, where V_{pi} and V_{qi} are the feature vectors of image P and query image Q respectively with size 'n'.

$$AED = \frac{1}{n} \sqrt{\sum_{i=1}^n (V_{pi} - V_{qi})^2} \quad (1)$$

Here total seven different image transforms are considered which can be listed as discrete cosine transform (DCT) [1,10,21,22,24], Walsh transform matrix [1,11,18,19,26], Haar Transform [28,29], Kekre transform [1,8,12,13,15,22], Discrete Sine Transform (DST) [36,37,40], Slant Transform [38, 39]. A discrete Hartley transform (DHT) [30-33]

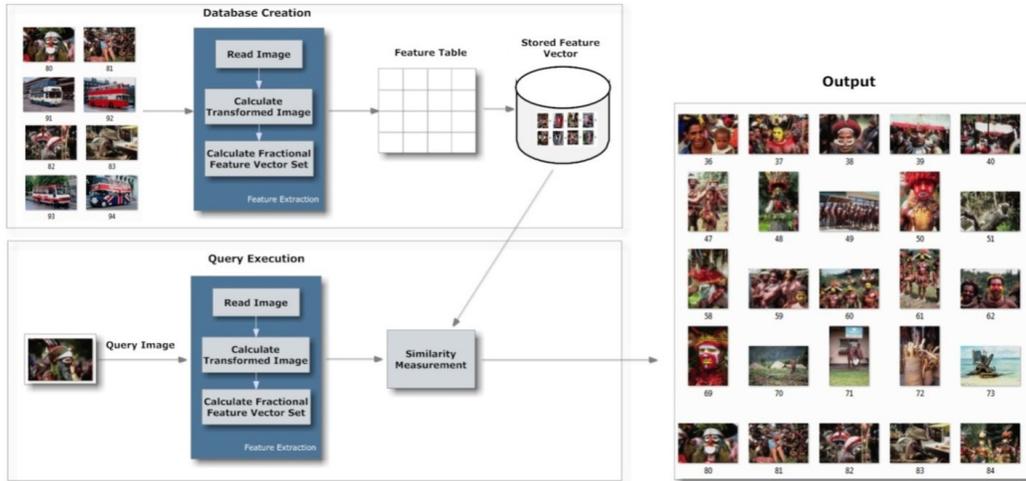


FIGURE 1.a. Flowchart of proposed CBIR Technique

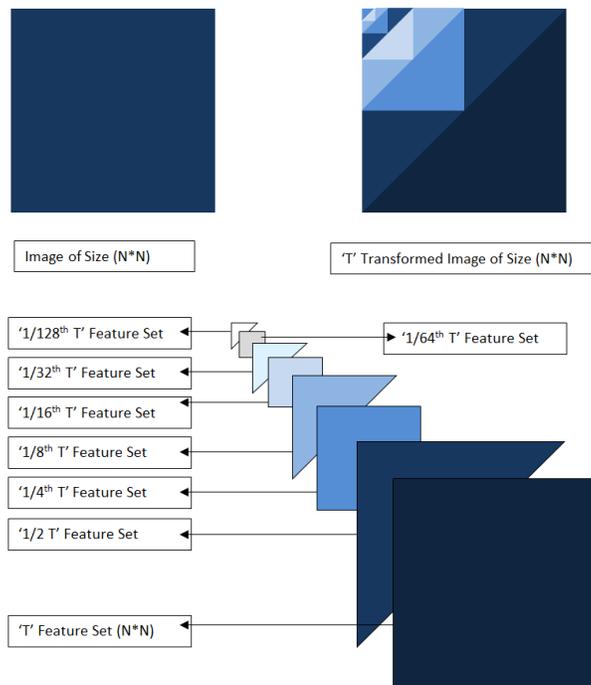


FIGURE 1b. Feature Extraction for Proposed CBIR Techniques

FIGURE 1. Proposed CBIR Techniques using fractional Coefficients of Transformed Images [43]

TABLE 1. Computational Complexity for applying transforms to image of size NxN [1,43]

Transform	Number of Additions	Number of Multiplications	Total Additions for transform of 256x256 image	Computations Comparison (For 256x256 image)
	For NxN image			
DCT	$2N^2(N-1)$	$N^2(2N)$	301858816	100
Walsh	$2N^2(N-1)$	0	33423360	11.07
Haar	$2N^2 \log_2(N)$	0	1048576	0.35
Kekre	$N[N(N+1)-2]$	$2N(N-2)$	17882624	5.92
DST	$2N^2(N-1)$	$N^2(2N)$	301858816	100
Slant	$2N^2(N-1)$	$N^2(N)$	167641088	55.54
Hartley	$2N^2(N-1)$	$N^2(2N)$	301858816	100

[Here one multiplication is considered as eight additions for second last row computations and DCT computations are considered to be 100% for comparison in last row]

2. PROPOSED CBIR-GRAY TECHNIQUES

The flowchart of proposed CBIR technique is given in figure 1.a for feature extraction and query execution. Figure 1.b explains the feature sets extraction used to extract feature sets for proposed CBIR techniques using fractional coefficients of transformed images.

2.1 Feature Extraction for feature vector 'T-Gray'

Here the feature vector space of the image of size $N \times N$ has $N \times N$ number of elements. This is obtained using following steps of T-Gray

- (i). Extract Red, Green and Blue components of the colour image.
- (ii). Take average of Red, Green and Blue components of respective pixels to get gray image.
- (iii). Apply the Transform 'T' on gray image to extract feature vector.
- (iv). The result is stored as the complete feature vector 'T-Gray' for the respective image.

Thus the feature vector database for DCT, Walsh, Haar, Kekre, DST, Slant, DHT transform are generated as DCT-Gray, Walsh-Gray, Haar-Gray, Kekre-Gray, DST-Gray, Slant-Gray, DHT-Gray respectively. Here the size of feature vector is $N \times N$ for every transform.

2.2 Feature Vector Database 'Fractional T-Gray'

The fractional coefficients of transformed image as shown in figure 1, are considered to form 'fractional T-Gray' feature vector databases. Here first 50% of coefficients from upper triangular part of feature vector 'T-Gray' are considered to prepare the feature vector database '50%-T-Gray' for every image as shown in figure 1. Thus DCT-Gray, Walsh-Gray, Haar-Gray, Kekre-Gray, DST-Gray, Slant-Gray, DHT-Gray feature databases are used to obtain new feature vector databases as 50%-DCT-Gray, 50%-Walsh-Gray, 50%-Haar-Gray, 50%-Kekre-Gray, 50%-DST-Gray, 50%-Slant-Gray, 50%-DHT-Gray respectively. Then per image first 25% number of coefficients (as shown in figure 1) form the feature vectors database DCT-Gray, Walsh-Gray, Haar-Gray, Kekre-Gray, DST-Gray, Slant-Gray, DHT-Gray are stored separately as feature vector databases as 25%-DCT-Gray, 25%-Walsh-Gray, 25%-Haar-Gray, 25%-Kekre-Gray, 25%-DST-Gray, 25%-Slant-Gray, 25%-DHT-Gray respectively. Then for each image in the database as shown in figure 1, fractional feature vector set for DCT-Gray, Walsh-Gray, Haar-Gray, Kekre-Gray, DST-Gray, Slant-Gray, DHT-Gray using 25%, 12.5%, 6.25%, 3.125%, 1.5625%, 0.7813%, 0.39%, 0.195%, 0.097%, 0.048%, 0.024%, 0.012% and 0.06% of total coefficients are formed.

2.3 Query Execution for 'T-Gray' CBIR

Here the feature set of size $N \times N$ for the query image is extracted using transform 'T'. This feature set is compared with each entry from the feature database using Euclidian distance as similarity measure. Thus DCT, Walsh, Haar, Kekre, DST, Slant, DHT transform based feature sets are extracted from query image and are compared respectively with DCT-Gray, Walsh-Gray, Haar-Gray, Kekre-Gray, DST-Gray, Slant-Gray, DHT-Gray feature sets using average Euclidian distance to find the best match in the database.

2.4 Query Execution for 'Fractional T-Gray' CBIR

For 50%-T-Gray query execution, only 50% number of coefficients of upper triangular part of 'T' transformed query image (with $N \times N$ coefficients) are considered for the CBIR and are compared with '50%-T-Gray' database feature set for Euclidian distance computations. Thus DCT, Walsh, Haar, Kekre, DST, Slant, DHT transform based feature sets are extracted from the query image and are compared respectively with 50%-DCT-Gray, 50%-Walsh-Gray, 50%-Haar-Gray, 50%-Kekre-Gray, 50%-DST-Gray, 50%-Slant-Gray, 50%-DHT-Gray feature sets to find average Euclidian distances. For 25%, 12.5%, 6.25%, 3.125%, 1.5625%, 0.7813%, 0.39%, 0.195%, 0.097%, 0.048%, 0.024%, 0.012% and 0.06% T-Gray based query execution, the feature set of the respective percentages are considered from the 'T' transformed $N \times N$ image as shown in figure 1, to be compared with the respective percentage T-Gray feature set database to find average Euclidian distance.

3. PROPOSED CBIR-RGB TECHNIQUES

3.1 Feature Extraction for feature vector ‘T-RGB’

Here the feature vector space of the image of size $N \times N \times 3$ has $N \times N \times 3$ number of elements. This is obtained using following steps of T-RGB

- (i). Extract Red, Green and Blue components of the color image.
- (ii). Apply the Transform ‘T’ on individual color planes of image to extract feature vector.
- (iii). The result is stored as the complete feature vector ‘T-RGB’ for the respective image.

Thus the feature vector database for DCT, Walsh, Haar, Kekre, DST, Slant, DHT transform is generated as DCT-RGB, Walsh-RGB, Haar-RGB, Kekre-RGB, DST-RGB, Slant-RGB, DHT-RGB respectively. Here the size of feature database is $N \times N \times 3$.

3.2 Query Execution for ‘T-RGB’ CBIR

Here the feature set of $N \times N \times 3$ for the query image is extracted using transform ‘T’ applied on the red, green and blue planes of query image. This feature set is compared with other feature sets in feature database using Euclidian distance as similarity measure. Thus DCT, Walsh, Haar, Kekre, DST, Slant, DHT transform based feature sets are extracted for query image and are compared respectively with DCT-RGB, Walsh-RGB, Haar-RGB, Kekre-RGB, DST-RGB, Slant-RGB, DHT-RGB feature sets to find Euclidian distance.

3.3 CBIR using ‘Fractional-T-RGB’

As explained in section 2.1 to 2.4 and section 3.1 – 3.3 the ‘T-RGB’ feature extraction and query execution are extended to get 50%,25%, 12.5%, 6.25%, 3.125%, 1.5625% ,0.7813%, 0.39%, 0.195%, 0.097%, 0.048%, 0.024%, 0.012% and 0.006% of T-RGB image retrieval techniques.

4. IMPLEMENTATION

The proposed CBIR methods are tested using a test bed of 1000 variable size images spread across 11 categories and taken from Wang image database [15]. The categories and distribution of the images is shown in table 2. Programming is done in MATLAB 7.0 using a computer with Intel Core 2 Duo Processor T8100 (2.1GHz) and 2 GB RAM.

TABLE 2. Image Database Category-Wise Distribution

Category	Monuments	Mountains	Beaches	Elephants	Roses	Tribes
No.of Images	99	61	99	99	99	85
Category	Horses	Dinosaurs	Airplanes	Buses	Sunrise	
No.of Images	99	99	100	99	61	



FIGURE 2. Sample Database Images
[Image database contains total 1000 images with 11 categories]

Figure 2 gives the sample database images from all categories of images considered in test bed image database. Precision and recall are used as statistical comparison parameters [1,2] for the proposed CBIR techniques. The standard definitions of these two measures are given by following equations.

$$Precision = \frac{Number_of_relevant_images_retrieved}{Total_number_of_images_retrieved} \tag{2}$$

$$Recall = \frac{Number_of_relevant_images_retrieved}{Total_number_of_relevant_images_in_database} \tag{3}$$

5. RESULTS AND DISCUSSION

The performance of each proposed CBIR technique is tested by firing 55 queries (5 from each category) per technique on the database of 1000 variable size generic images spread across 11 categories. Average Euclidian distance is used as similarity measure. The average precision and average recall are computed by grouping the number of retrieved images sorted according to ascending average Euclidian distances with the query image. In all transforms, the average precision and average recall values for CBIR using fractional coefficients are higher than CBIR using full set of coefficients. The CBIR-Gray techniques are giving higher values of crossover points than CBIR-Gray techniques indicating better performance. The crossover point of precision and recall of the CBIR techniques acts as one of the important parameters to judge their performance [1,2,19,20].

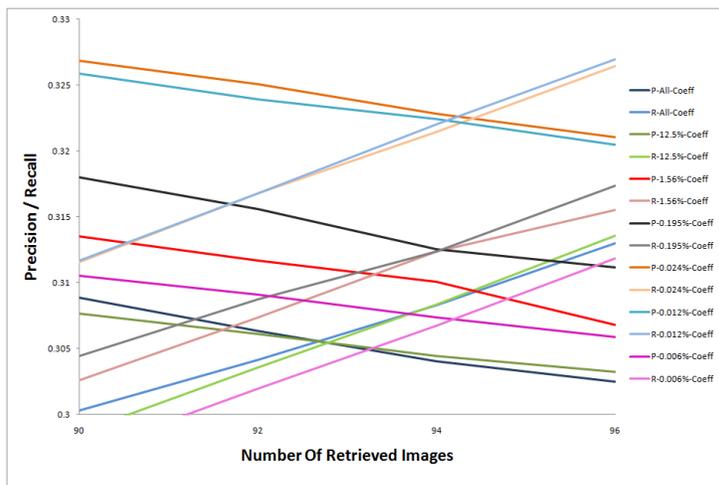


FIGURE 3.a. DCT-Gray based CBIR

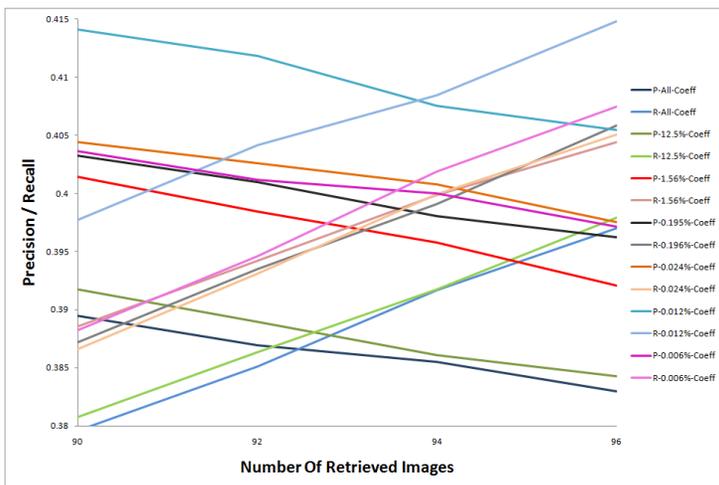


FIGURE 3.b. DCT-RGB based CBIR

FIGURE 3. Crossover Point of Precision and Recall for DCT based CBIR.

Figure 3 shows the precision-recall crossover points plotted against number of retrieved images for proposed image retrieval techniques using DCT. Uniformly in all image retrieval techniques based on gray DCT and color DCT features 0.012% fractional feature set ($1/8192^{\text{th}}$ of total coefficients) based image retrieval gives highest precision and recall values. Figure 3.a gives average precision/recall values plotted against number of retrieved images for all DCT-Gray image retrieval techniques. Precision/recall values for DCT-RGB image retrieval techniques are plotted in figure 3.b.

Figures 4.a and 4.b respectively shows the graphs of precision/recall values plotted against number of retrieved images for Walsh-Gray and Walsh-RGB based image retrieval techniques. Here $1/4096^{\text{th}}$ fractional coefficients (0.024% of total Walsh transformed coefficients) based image retrieval gives the highest precision/recall crossover values.

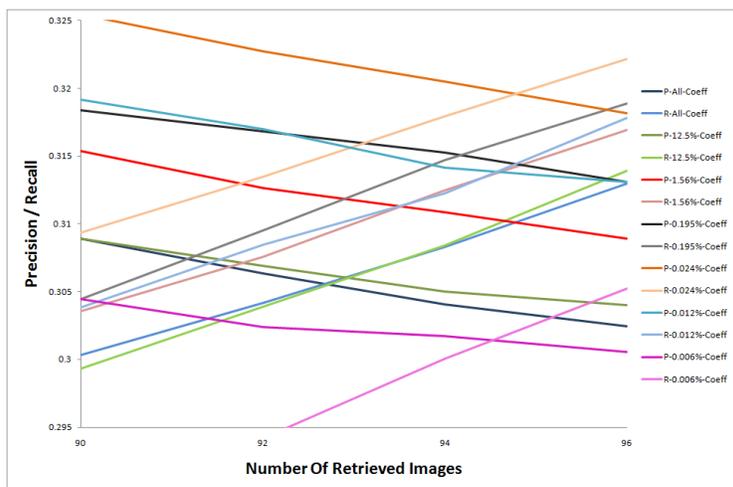


FIGURE 4.a. Walsh-Gray based CBIR

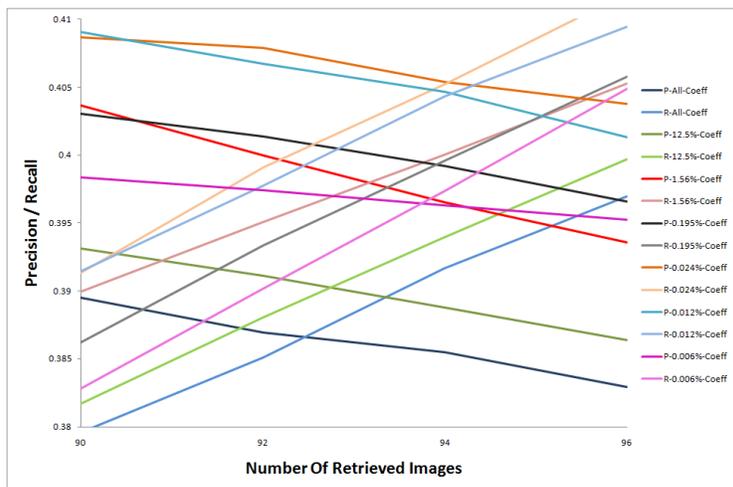


FIGURE 4.b. Walsh-RGB based CBIR

FIGURE 4. Crossover Point of Precision and Recall for Walsh Transform based CBIR

Figure 5 shows the precision-recall crossover points plotted against number of retrieved images for proposed image retrieval techniques using Haar Transform. Uniformly in all image retrieval techniques based on gray Haar and color Haar features 0.024% fractional feature set ($1/4096^{\text{th}}$ of total coefficients) based image retrieval gives highest precision and recall values. Figures 5.a and 5.b give average precision/recall values plotted against number of retrieved

images for all Haar-Gray and Haar-RGB image retrieval techniques respectively. Figure 6.a gives average precision/recall values plotted against number of retrieved images for all Kekre-Gray image retrieval techniques. Precision/recall values for Kekre-RGB image retrieval techniques are plotted in figure 6.b. Here 1/32th fractional coefficients (3.125% of total Kekre transformed coefficients) based image retrieval gives the highest precision/recall crossover values specifying the best performance.

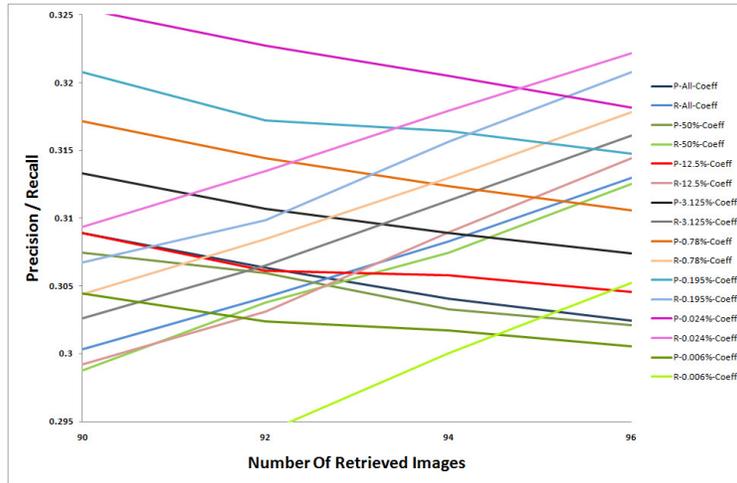


FIGURE 5.a. Haar-Gray based CBIR

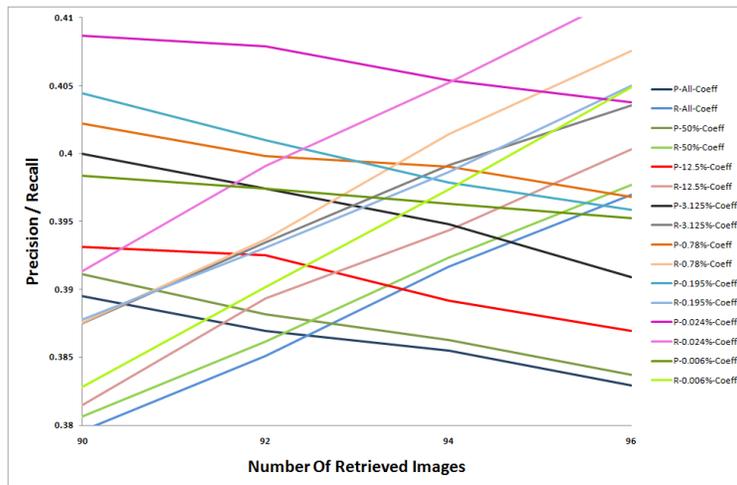


FIGURE 5.b. Haar-RGB based CBIR

FIGURE 5. Crossover Point of Precision and Recall for Haar Transform based CBIR

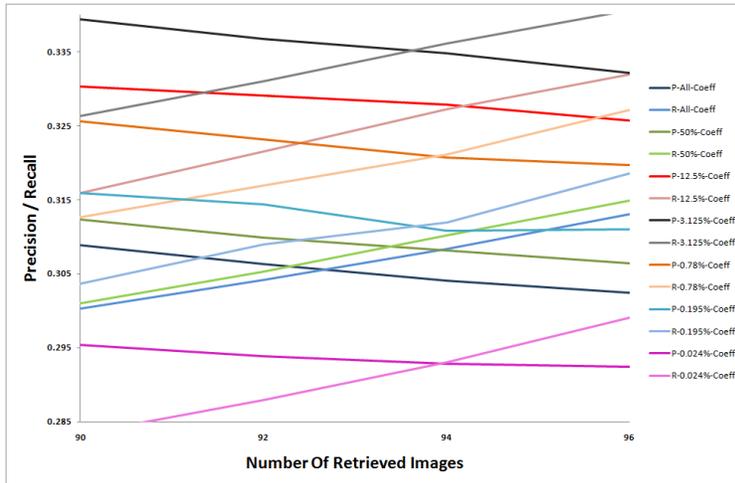


FIGURE 6.a. Kekre-Gray based CBIR

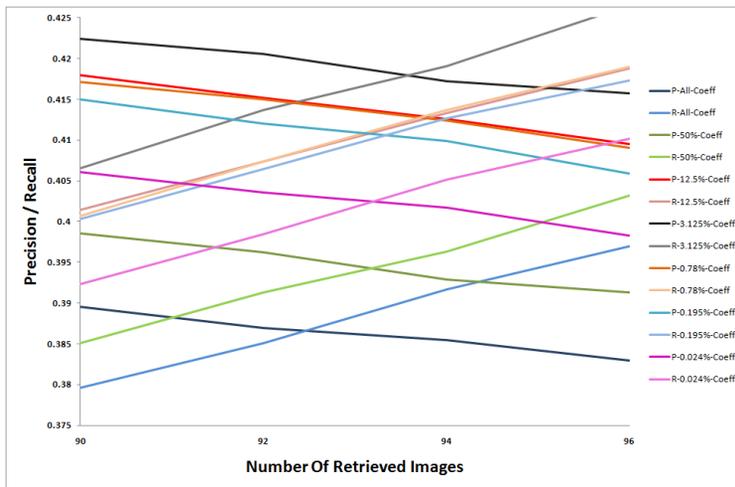


FIGURE 6.b. Kekre-RGB based CBIR

FIGURE 6. Crossover Point of Precision and Recall for Kekre Transform based CBIR

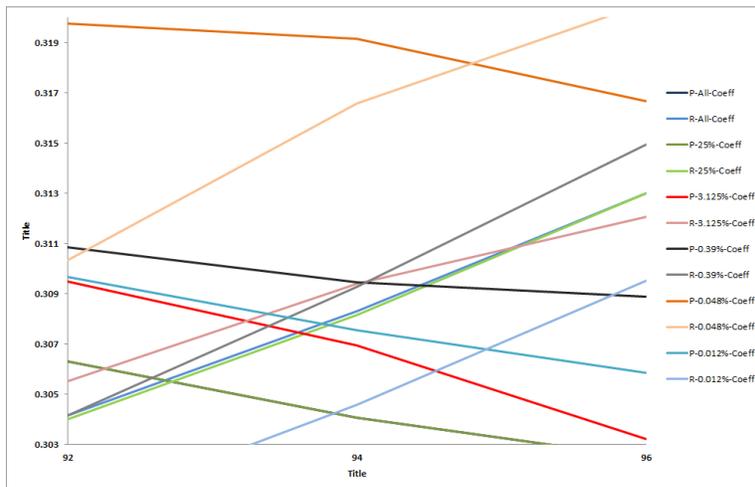


FIGURE 7.a. DST-Gray based CBIR

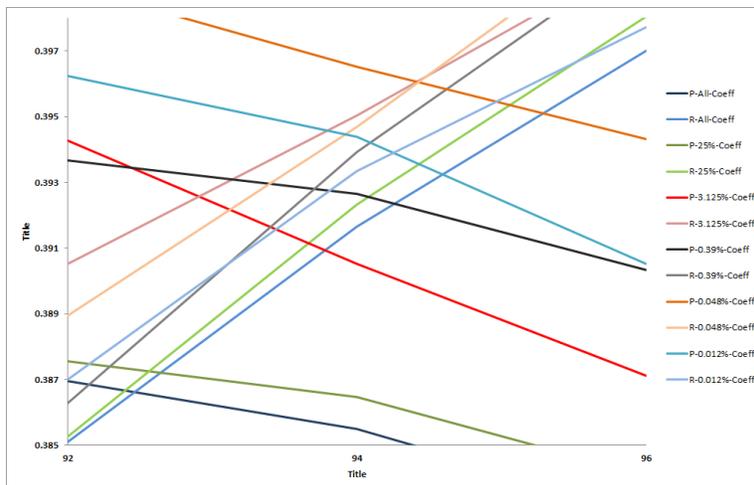


FIGURE 7.b. DST-RGB based CBIR

FIGURE 7. Crossover Point of Precision and Recall for DST based CBIR

Figure 7.a gives average precision/recall values plotted against number of retrieved images for all DST-Gray image retrieval techniques. Precision/recall values for DST-RGB image retrieval techniques are plotted in figure 7.b. Here 1/2048th fractional coefficients (0.048% of total DST transformed coefficients) based image retrieval gives the highest precision/recall crossover values specifying the best performance when using Discrete Sine Transform.

Figure 8.a gives average precision/recall values plotted against number of retrieved images for all Slant-Gray image retrieval techniques. Precision/recall values for Slant-Gray image retrieval techniques are plotted in figure 8.b. Here 1/4th fractional coefficients (25% of total Slant transformed coefficients) based image retrieval gives the highest precision/recall crossover values specifying the best performance when using Slant Transform on a gray image and 1/8th fractional coefficients (12.5% of total Slant transformed coefficients) based image retrieval gives the highest precision/recall crossover values specifying the best performance when using Slant Transform on a color image.

Figure 9.a gives average precision/recall values plotted against number of retrieved images for all DHT-Gray image retrieval techniques. Precision/recall values for DHT-Gray image retrieval techniques are plotted in figure 9.b. Here 1/2 fractional coefficients (0.50% of total DHT transformed coefficients) based image retrieval gives the highest precision/recall crossover values specifying the best performance when using Discrete Hartley Transform.

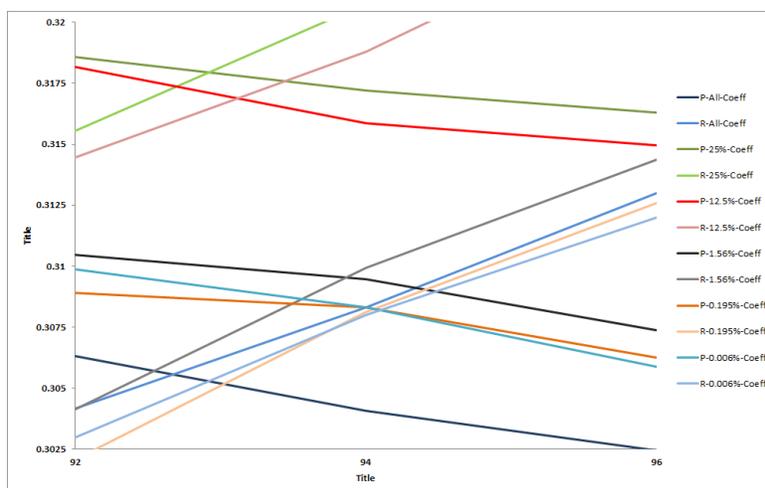


FIGURE 8.a. Slant-Gray based CBIR

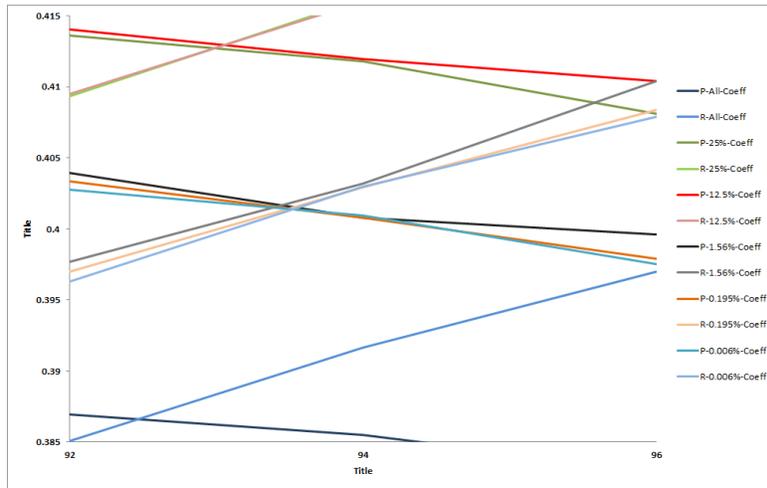


FIGURE 8.b. Slant-RGB based CBIR

FIGURE 8. Crossover Point of Precision and Recall for Slant Transform based CBIR

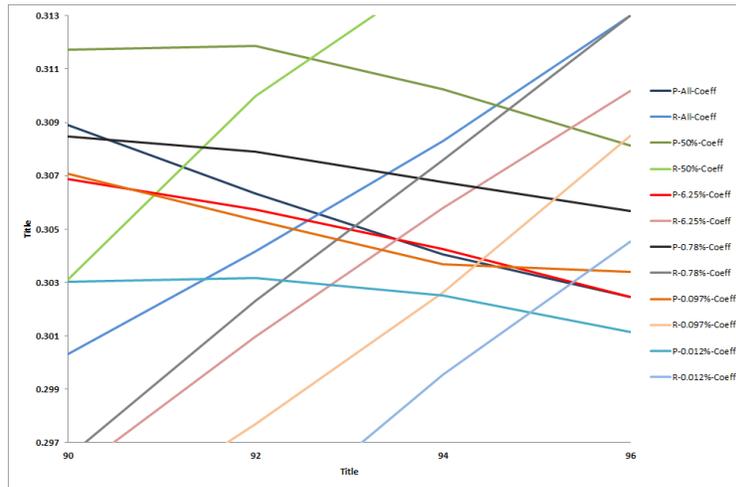


FIGURE 9.a. DHT-Gray based CBIR

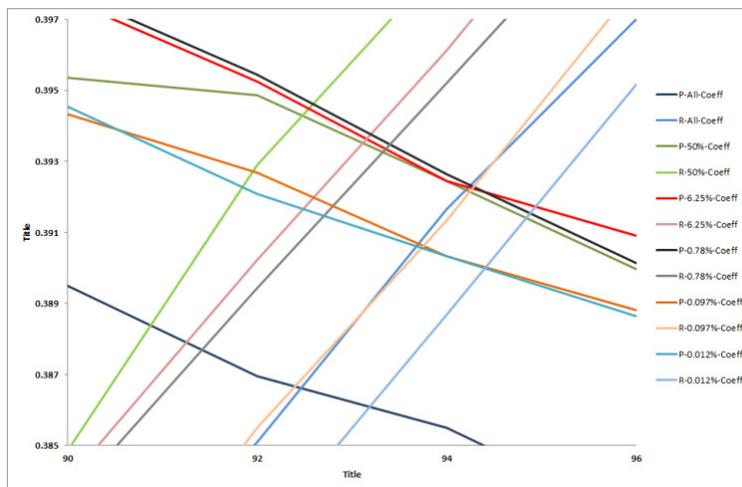


FIGURE 9.b. DHT-RGB based CBIR

FIGURE 9. Crossover Point of Precision and Recall for DHT based CBIR

Figure 10 shows the performance comparison of all the seven transforms for proposed CBIR techniques. Figure 10.a is indicating the crossover points of DCT-Gray, Walsh-Gray, Haar-Gray, Kekre-Gray, DST-Gray, Slant-Gray, and DHT-Gray CBIR for all considered feature vectors (percentage of coefficients of transformed gray images). Here for upto 25% of coefficients Slant transform performs better than all discussed transforms then till 0.78 % of coefficients Kekre transform performs better than all discussed transforms after which Haar transform outperforms other transforms up to 0.012 % of coefficients and finally for 0.006 % of coefficients DCT gives highest crossover point value, as energy compaction in Haar and DCT transform is better than other discussed transforms. For Kekre-Gray CBIR the performance improves with decreasing feature vector size from 100% to 0.195% and then drops indicating 0.195% as best fractional coefficients. In DCT-Gray CBIR the performance is improved till 0.012% and then drops. Overall in all, CBIR using Kekre transform with 3.125 % of fractional coefficients gives the best performance for Gray-CBIR techniques discussed here. Figure 10.b indicates the performance comparison of DCT-RGB, Walsh- RGB, Haar- RGB, Kekre- RGB, DST-Gray, Slant-Gray, and DHT-Gray CBIR with different percentage of fractional coefficients. Here Slant-RGB CBIR outperforms all other transforms till 12.5% of coefficients as feature vector then Kekre-RGB CBIR outperforms all other transforms till 0.097% of coefficients as feature vector then Walsh-RGB CBIR takes over till 0.024% then DCT-RGB performs best for 0.012% of coefficients. In Walsh-RGB and Haar-RGB CBIR the feature vector with 0.024% of coefficients gives best performance, in DCT-RGB CBIR 0.012% of coefficients shows highest crossover value of average precision and average recall and Kekre transform gives the best performance when 6.25% of coefficients are considered. In all, CBIR using Kekre transform with 6.25 % of fractional coefficients gives the best performance for RGB-CBIR techniques discussed here.

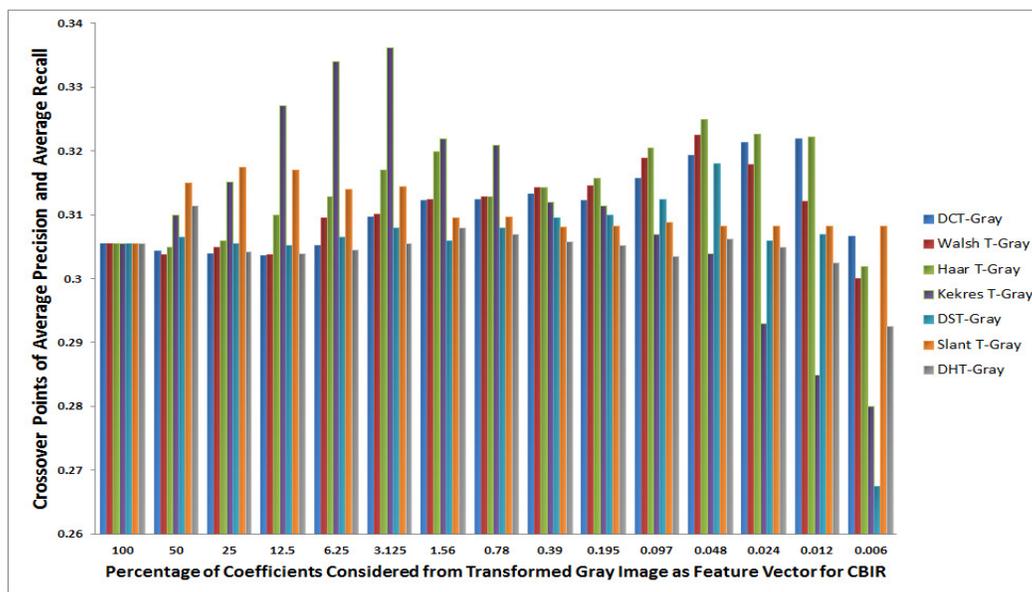


FIGURE 10.a. Transform Comparison in Gray based CBIR

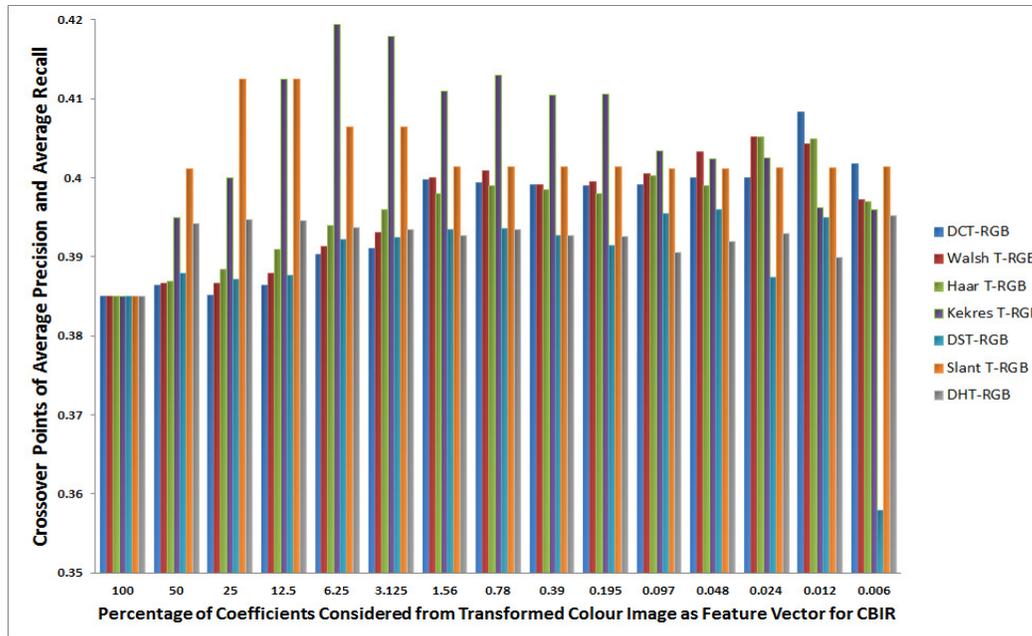


Fig.10.b. Transform Comparison in Color based CBIR

FIGURE 10. Performance Comparison Using Fractional Coefficients across all discussed Transforms

6. CONCLUSION

In the information age where the size of image databases is growing exponentially more precise retrieval techniques are needed, for finding relatively similar images. Computational complexity and retrieval efficiency are the key objectives in the image retrieval system. Nevertheless it is very difficult to reduce the computations and improve the performance of image retrieval technique.

Here the performance of image retrieval is improved using fractional coefficients of transformed images at reduced computational complexity. Fairly large numbers of popular image transforms are considered here with newly introduced Kekre transform. In all transforms (DCT, Walsh, Haar, Kekre, DST, Slant and DHT), the average precision and average recall values for CBIR using fractional coefficients are higher than CBIR using full set of coefficients. Hence the feature vector size for image retrieval could be greatly reduced, which ultimately will result in faster query execution in CBIR with better performance. In all Kekre transform with fractional coefficients (3.125 % in Gray and 6.25 % in RGB) gives best performance with highest crossover points of average precision and average recall. Feature extraction using Kekre transform is also computationally lighter as compared to DCT or Walsh transform. Thus feature extraction in lesser time is possible with increased performance.

Finally the conclusion that the fractional coefficients gives better discrimination capability in CBIR than the complete set of transformed coefficients and image retrieval with better performance at much faster rate can be done from the proposed techniques and experimentation done.

REFERENCES

- [1] H.B.Kekre, Sudeep D. Thepade, "Improving the Performance of Image Retrieval using Partial Coefficients of Transformed Image", International Journal of Information Retrieval (IJIR), Serials Publications, Volume 2, Issue 1, 2009, pp. 72-79(ISSN: 0974-6285)
- [2] 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. ConceicaoRodrigo College of Engg., Mumbai. Is uploaded on online ACM portal.

- [3] H.B.Kekre, Sudeep D. Thepade, "Scaling Invariant Fusion of Image Pieces in Panorama Making and Novel Image Blending Technique", International Journal on Imaging (IJI), www.ceser.res.in/iji.html, Volume 1, No. A08, pp. 31-46, Autumn 2008.
- [4] Hirata K. and Kato T. "Query by visual example – content-based image retrieval", In Proc. of Third International Conference on Extending Database Technology, EDBT'92, 1992, pp 56-71
- [5] H.B.Kekre, Sudeep D. Thepade, "Rendering Futuristic Image Retrieval System", National Conference on Enhancements in Computer, Comm. and Information Technology, EC2IT-2009, 20-21 Mar 2009, K.J.Somaiya COE, Vidyavihar, Mumbai-77.
- [6] Minh N. Do, MartinVetterli, "Wavelet-Based Texture Retrieval Using Generalized Gaussian Density and Kullback-Leibler Distance", IEEE Transactions On Image Processing, Volume 11, Number 2, pp.146-158, February 2002.
- [7] B.G.Prasad, K.K. Biswas, and S. K. Gupta, "Region –based image retrieval using integrated color, shape, and location index", International Journal on Computer Vision and Image Understanding Special Issue: Colour for Image Indexing and Retrieval, Volume 94, Issues 1-3, April-June 2004, pp.193-233.
- [8] H.B.Kekre, Sudeep D. Thepade, "Creating the Color Panoramic View using Medley of Grayscale and Color Partial Images ", WASET International Journal of Electrical, Computer and System Engineering (IJCSE), Volume 2, No. 3, Summer 2008. Available online at www.waset.org/ijecse/v2/v2-3-26.pdf.
- [9] StianEdvardsen, "Classification of Images using color, CBIR Distance Measures and Genetic Programming", Ph.D. Thesis, Master of science in Informatics, Norwegian university of science and Technology, Dept. of computer and Information science, June 2006.
- [10] H.B.Kekre, TanujaSarode, Sudeep D. Thepade, "DCT Applied to Row Mean and Column Vectors in Fingerprint Identification", In Proceedings of International Conference on Computer Networks and Security (ICCNS), 27-28 Sept. 2008, VIT, Pune.
- [11] Zhibin Pan, Kotani K., Ohmi T., "Enhanced fast encoding method for vector quantization by finding an optimally-ordered Walsh transform kernel", ICIP 2005, IEEE International Conference, Volume 1, pp I - 573-6, Sept. 2005.
- [12] H.B.kekre, Sudeep D. Thepade, "Improving 'Color to Gray and Back' using Kekre LUV Color Space", IEEE Int. Advanced Computing Conference 2009 (IACC'09), Thapar University, Patiala, INDIA, 6-7 March 2009. Is uploaded at IEEE Xplore.
- [13] H.B.Kekre, Sudeep D. Thepade, "Image Blending in Vista Creation using Kekre LUV Color Space", SPIT-IEEE Colloquium and International Conference, Sardar Patel Institute of Technology, Andheri, Mumbai, 04-05 Feb 2008.
- [14] H.B.Kekre, Sudeep D. Thepade, "Color Traits Transfer to Grayscale Images", In Proc.of IEEE First International Conference on Emerging Trends in Engg. & Technology, (ICETET-08), G.H.Raisoni COE, Nagpur, INDIA. Uploaded on online IEEE Xplore.
- [15] <http://wang.ist.psu.edu/docs/related/Image.org> (Last referred on 23 Sept 2008)
- [16] H.B.Kekre, Sudeep D. Thepade, "Using YUV Color Space to Hoist the Performance of Block Truncation Coding for Image Retrieval", IEEE Int. Advanced Computing Conference 2009 (IACC'09), Thapar University, Patiala, INDIA, 6-7 March 2009.
- [17] H.B.Kekre, Sudeep D. Thepade, ArchanaAthawale, Anant Shah, PrathmeshVerlekar, Suraj Shirke, "Energy Compaction and Image Splitting for Image Retrieval using Kekre Transform over Row and Column Feature Vectors", International Journal of Computer Science and Network Security (IJSNS), Volume:10, Number 1, January 2010, (ISSN: 1738-7906) Available at www.IJSNS.org.
- [18] H.B.Kekre, Sudeep D. Thepade, ArchanaAthawale, Anant Shah, PrathmeshVerlekar, SurajShirke, "Walsh Transform over Row Mean and Column Mean using Image Fragmentation and Energy Compaction for Image Retrieval", International Journal on Computer Science and Engineering (IJCSE), Volume 2S, Issue1, January 2010, (ISSN: 0975-3397). Available online at www.enggjournals.com/ijcse.

- [19] H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Color-Texture Features Extracted from Walshlet Pyramid", ICGST International Journal on Graphics, Vision and Image Processing (GVIP), Volume 10, Issue 1, Feb.2010, pp.9-18, Available online www.icgst.com/gvip/Volume10/Issue1/P1150938876.html
- [20] H.B.Kekre, Sudeep D. Thepade, "Color Based Image Retrieval using Amendment Block Truncation Coding with YCbCr Color Space", International Journal on Imaging (IJI), Vol. 2, No. A09, Autumn 2009, pp. 2-14. Available at www.ceser.res.in/iji.html.
- [21] H.B.Kekre, TanujaSarode, Sudeep D. Thepade, "Color-Texture Feature based Image Retrieval using DCT applied on Kekre Median Codebook", International Journal on Imaging (IJI), Vol. 2, No. A09, Autumn 2009, pp. 55-65. Available at www.ceser.res.in/iji.html.
- [22] H.B.Kekre, Sudeep D. Thepade, "Image Retrieval using Non-Involutorial Orthogonal Kekre Transform", International Journal of Multidisciplinary Research and Advances in Engineering (IJMRAE), Ascent Publication House, 2009, Volume 1, No.1, pp 189-203, 2009. Abstract available online at www.ascent-journals.com (ISSN: 0975-7074)
- [23] H.B.Kekre, Sudeep D. Thepade, "Boosting Block Truncation Coding using Kekre LUV Color Space for Image Retrieval", WASET International Journal of Electrical, Computer and System Engineering (IJECSE), Volume 2, Number 3, pp. 172-180, Summer 2008. Available online at <http://www.waset.org/ijecse/v2/v2-3-23.pdf>
- [24] H.B.Kekre, Sudeep D. Thepade, ArchanaAthawale, Anant Shah, PrathmeshVerlekar, SurajShirke, "Performance Evaluation of Image Retrieval using Energy Compaction and Image Tiling over DCT Row Mean and DCT Column Mean", Springer-International Conference on Contours of Computing Technology (Thinkquest-2010), BabasahebGawde Institute of Technology, Mumbai, 13-14 March 2010, The paper will be uploaded on online Springerlink.
- [25] H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, VaishaliSuryavanshi, "Improved Texture Feature Based Image Retrieval using Kekre Fast Codebook Generation Algorithm", Springer-Int. Conference on Contours of Computing Technology (Thinkquest-2010), BabasahebGawde Institute of Technology, Mumbai, 13-14 March 2010, The paper will be uploaded on online Springerlink.
- [26] H.B.Kekre, Tanuja K. Sarode, Sudeep D. Thepade, "Image Retrieval by Kekre Transform Applied on Each Row of Walsh Transformed VQ Codebook", (Invited), ACM-Int. Conference and Workshop on Emerging Trends in Technology (ICWET 2010),Thakur College of Engg. & Tech., Mumbai, 26-27 Feb 2010, The paper is uploaded at ACM Portal.
- [27] H.B.Kekre, Sudeep D. Thepade, AkshayMaloo, "Image Retrieval using Fractional Coefficients of Transformed Image using DCT and Walsh Transform", International Journal of Engineering Science and Technology (IJEST), Volume 2, Number 4, 2010, pp.362-371.
- [28] Haar, Alfred, "ZurTheorie der orthogonalen Funktionen systeme". (German), MathematischeAnnalen, volume 69, No. 3, 1910, pp. 331-371.
- [29] Charles K. Chui, "An Introduction to Wavelets", Academic Press, 1992, San Diego, ISBN 0585470901.
- [30] R. N. Bracewell, "Discrete Hartley transform," Journal of Opt. Soc. America, Volume 73, Number 12, pp. 1832-183 , 1983.
- [31] R. N. Bracewell, "The fast Hartley transform," Proc. of IEEE Vol. 72, Num. 8, pp.1010-1018 ,1984.
- [32] R. N. Bracewell, 'The Hartley Transform", Oxford Univ. Press, New York, 1986.
- [33] R. N. Bracewell, "Computing with the Hartley Transform," Computers in Physics 9 (4), 373-379 (1995).
- [34] R. V. L. Hartley, "A more symmetrical Fourier analysis applied to transmission problems," Proc. IRE 30, 144-150 (1942).
- [35] H. V. Sorensen, D. L. Jones, M. T. Heideman, and C. S. Burrus, "Real-valued fast Fourier transform algorithms," IEEE Trans. Acoust. Speech Sig. Processing ASSP-35, pp.849-863, 1987.
- [36] S. A. Martucci, "Symmetric convolution and the discrete sine and cosine transforms," IEEE Trans. Sig. Processing SP-42, pp. 1038-1051, 1994.

- [37] Matteo Frigo and Steven G. Johnson: FFTW, <http://www.fftw.org/>. A free (GPL) C library that can compute fast DSTs (types I-IV) in one or more dimensions, of arbitrary size. Also M. Frigo and S. G. Johnson, "The Design and Implementation of FFTW3," Proceedings of the IEEE Volume 93, Number 2, pp.216–231, 2005.
- [38] Rant W K, Welch L R and Chen W H, "Slant Transforms for Image Coding". Proc. Symp. Applications of Walsh Functions: 229-434, 1972.
- [39] Zheng-Xin Hou, Ni-Ni Xu, Hong Chen, Xue-Lei Li, "Fast Slant Transform with Sequence Increment and Its Applications in Image Compression", In Proc. of the Third International Conference on Machine Learning and Cybernetics, Shanghai, 26-29 August 2004.
- [40] H.B.Kekre, J.K.Solanki, "Comparative performance of various trigonometric unitary transforms for transform image coding", International Journal of Electronics, 1362-3060, Volume 44, Issue 3, 1978, pp. 305 – 315.
- [41] Ch.SrinivasaRao, S. Srinivas Kumar, B.N.Chatterji, "Content Based Image Retrieval using Contourlet Transform", ICGST-GVIP Journal, Volume 7, Issue 3, November 2007, pp.9-15.
- [42] Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, Ramesh Jain, "Content-Based Image Retrieval at the End of the Early Years", IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 22, Number 12, December 2000.
- [43] H.B.Kekre, Sudeep D. Thepade, Akshay Maloo, "Performance Comparison of Image Retrieval Using Fractional Coefficients of Transformed Image Using DCT, Walsh, Haar and Kekre Transform", CSC International Journal of Image Processing (CSC-IJIP), Volume 4, Issue 2, pp 142-155.

Frequency Domain Blockiness and Blurriness Meter for Image Quality Assessment

Qadri, Muhammad Tahir

*School of Computer Science and Electronic Engineering
University of Essex
Colchester, CO4 3SQ, UK*

mtqadr@essex.ac.uk

Tan, K.T.

*School of Electrical and Electronic Engineering
Singapore Polytechnic
Singapore*

kttan@sp.edu.sg

Ghanbari, Muhammad

*School of Computer Science and Electronic Engineering
University of Essex
Colchester, CO4 3SQ, UK*

ghan@essex.ac.uk

Abstract

Image and video compression introduces distortions (artefacts) to the coded image. The most prominent artefacts added are blockiness and blurriness. Many existing quality meters are normally distortion-specific. This paper proposes an objective quality meter for quantifying the combined blockiness and blurriness distortions in frequency domain. The model first applies edge detection and cancellation, then spatial masking to mimic the characteristics of the human visual system. Blockiness is then estimated by transforming image into frequency domain, followed by finding the ratio of harmonics to other AC components. Blurriness is determined by comparing the high frequency coefficients of the reference and coded images due to the fact that blurriness reduces the high frequency coefficients. Then, both blockiness and blurriness distortions are combined for a single quality metric. The meter is tested on blocky and blurred images from the LIVE image database, with a correlation coefficient of 95-96%.

Keywords: Blockiness and Blurriness Measurement, Full Reference, Image Quality Assessment

1. INTRODUCTION

Block-based digital image and video compression could lead to visible distortions in the coded image or video, with dominant artifacts like blockiness and blurriness. Over the years, enormous amount of efforts had been put, and many objective measurement models have been proposed. Many quality meters are distortion specific, for example, blockiness [1-3], blurriness [4-6] and ringing [7-8]. The most popular and widely used Quality Meter is the Peak Signal to Noise Ratio (PSNR) but it doesn't correlate well with the subjective assessment results. A reliable Image Quality Meter (IQM) should include all of the degradation artifacts (such as blockiness and blurriness) and weight these artifacts individually to combine into one IQM metric. In this paper, we demonstrate the use of frequency domain analysis to measure the blockiness and blurriness artifacts. Both distortions degrade the image quality; blockiness results in discontinuities at the block boundaries and blurriness reduces the sharpness of images. Since the blockiness generates periodic pattern, and blurring degrades the image quality due to loss of high spatial frequency components, it is possible to analyze these artifacts in the frequency domain instead of spatial domain. To quantify the image quality accurately, the Human Visual System (HVS) plays an important role. It is well known that the sensitivity of human eyes varies with the frequency change and therefore the presence of textual details at locations will tend to reduce the visibility

of distortion in that area. Therefore efficient masking of distortions according to the local spatial activity can lead to a better quality meter.

The rest of this paper is organized as follows: Section II gives an overview of Harmonics Analysis; Section III explains the design of the Image Quality meter; and Section IV shows the obtained results and section V lists the used references.

2. An Overview of Harmonic Amplitude Analysis

Blockiness and blurriness are the two main distortions in the image during coding process. Blockiness is due to luminance discontinuities across DCT block boundaries. As the DCT block size is fixed (or varies in the ratio of 2), the luminance discontinuity is periodic. If these luminance discontinuities are extracted and transformed into frequency domain, harmonics at certain spatial frequencies are created due to periodicity of the luminance discontinuities. The strength of these harmonics is proportional to the degree of blockiness in the coded picture [9]. Figure 1 shows the basic concept of Harmonics Amplitude Analysis (HAA) in frequency domain.

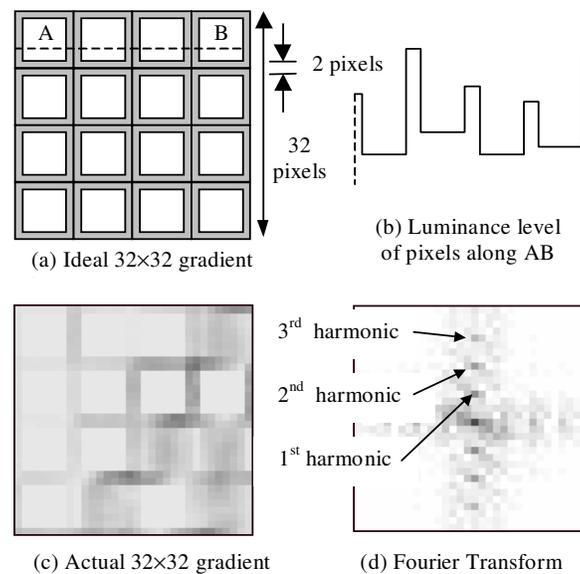


FIGURE 1: Basic concept of Harmonic Amplitude Analysis

Figure 1(a) is an ideal blocky gradient image of 32x32 while Figure 1(b) shows the luminance level transitions across 1 row of pixels. Figure 1(c) shows a 32x32 block extracted from the gradient image of a blocky picture, and Figure 1(d) is its Fourier transform, with the three vertical harmonics highlighted.

For the calculation of blockiness or blurriness index, the image is divided into blocks for block processing. The reason for block processing is that if we apply the FFT on whole image without block processing, then the chances of error is very high because the distortion might not be consistent and equal in every part of the image so the distortion is computed for each block locally and will be accumulated in the end as a single quality metric. The size 32x32 is chosen because the block size should be multiple of 8 (as DCT block size is 8x8) and the harmonics must have some distance among them to be recognized as harmonics that is why 32x32 window size is selected.

After dividing the image into blocks of 32x32 pixels, each block is transformed into frequency domain. As the frequency domain has two parts; the amplitude and phase, only the amplitude part is used in this work because blockiness or blurriness affects the amplitude part only.

For every block of 32x32, there will be 1 DC component and 32 AC components (2 repetitive groups of 16). There will be three harmonics for 32x32 window, i.e. H4, H8 and H12. For determining the harmonic frequency, recall the equation,

$$f_k = k * \left(\frac{W}{H}\right), \text{ for } k = 1, 2, \dots, \left(\frac{W}{H} - 1\right) \quad (1)$$

where, f_k is the frequency of harmonics in cycles per window (cpw) and as it also depends on the width of the FFT window w . The above equation represents the harmonics in frequency domain. The next section explains the design of full reference quality meter mainly focusing the combination of two distortions.

3. DESIGN OF THE IMAGE QUALITY METER

The main emphasis of this paper is to develop a distortion meter with combined blockiness and blurriness distortions. Only blockiness meter is not good in estimating quality of lightly compressed image (very little blockiness) therefore adding a blur meter will help to compensate for that weakness. In the following parts the combined blockiness and blurriness quality meter which is designed for full reference (FR) mode is explained. It consists of 3 main parts; 1) blockiness estimation; 2) blurriness estimation and 3) combining the two distortions.

3.1 Blockiness Estimation

The blockiness estimation stage consists of 4 stages; Edge Detection; Edge Cancellation; Spatial Masking; and Blockiness Estimation in the frequency domain. Figure 2 shows a block diagram of the meter (blockiness part).

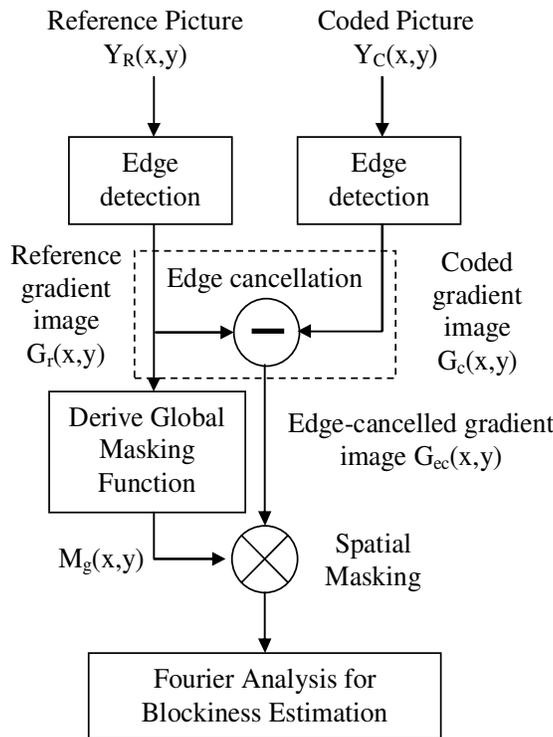


FIGURE 2: Block Diagram of the blockiness detector

3.1.1 Edge Detection

Edge detection is used to determine the sharp luminance edges from the reference image. These sharp luminance edges are either due to the blockiness artifact introduced in coding process or due to the textual details present in reference image. This spatial activity of both, reference and

coded images, are determined by using sobel edge detectors. The edge detection is performed horizontally and then vertically on both images.

3.1.2 Edge Cancellation

Some heavy textual details in the reference image might also be misinterpreted as blockiness, so the edge cancellation process is performed between edge detected versions of coded and reference images to reduce the chance of misinterpreting the contextual details as blockiness. It leaves only the edges due to blockiness. Also note that, the coding process also removes the high frequency coefficients so the simple subtraction of reference image from coded image will not perform effective edge cancellation. So, the pixel comparison is done between the coded and reference image to check if the pixel value of coded image is greater than the reference pixel value only then subtract the reference picture from the coded image as explained in equation below. The equation below is used to apply the edge cancellation process.

$$G_{ec}(x,y) = \begin{cases} G_c(x,y) - G_r(x,y) & \text{if } G_c(x,y) > G_r(x,y) \\ 0 & \text{Else} \end{cases}$$

where, $G_{ec}(x,y)$ is the edge cancelled gradient image, $G_c(x,y)$ is the edge detected coded image and $G_r(x,y)$ is the edge detected reference image. Now, the edge cancelled gradient image contains spatial activity only due to blocking artifact.

3.1.3 Spatial Masking

The next part is to mask the edge cancelled gradient image according to the spatial activity present in the reference image. The spatial masking is performed because perception of blockiness in detailed areas can be more masked than in low detailed areas of the picture. To derive the global masking function $M_g(x,y)$, the edge detected version of the reference image is used. For each pixel in the gradient reference image $G_r(x,y)$, a local masking function $m_g(x,y,d)$ is derived. Then these local masking functions are used to obtain the global masking function $M_g(x,y)$ which contains the spatial activity of the reference image. In this work, the spatial masking from Fiorentini and Zoli [10] is adopted. The masked reference image is convolved with the edge cancelled gradient image. After the edge cancellation and spatial masking, the edge cancelled spatial masked gradient image is ready for Fourier transformation for quantifying the blockiness.

3.1.4 Frequency Domain Analysis

Based on the concept of harmonics in frequency domain, the Harmonic Amplitude Analysis (HAA) method is proposed for blockiness estimation, in which the comparison of the strength of harmonics with other AC components is performed. Since blockiness is the abrupt luminance changes so it creates the harmonics in frequency domain and the strength of these harmonics are directly proportional to the amount of blockiness in the coded image.

The local blockiness index for each block is calculated by comparing the strength of harmonics with the rest of AC components. The higher the blockiness in an image, the higher will be the harmonics to AC components ratio and higher will be the blockiness. The equation for the Amplitude Harmonic Analysis ratio is given below:

$$R = \frac{\sum(H_4 + H_8 + H_{12})}{\sum_{n=1}^{15} H_n}$$

The amplitude harmonic analysis ratio R is calculated both vertically and horizontally and then accumulated in the end for a single blockiness metric.

3.2 Blurriness Estimation

Blurriness occurs over the whole image by reducing the sharpness of image due to the loss of high frequency coefficients and it is easier to determine blurriness in frequency domain by analyzing the high frequency coefficients. The amount of blurriness is estimated by comparing the high frequency coefficients of reference and coded images.

The blurriness distortion also needs to be masked according to the local spatial activity from the reference image. The same spatial masking concept (which was used for blockiness meter) is used to determine the spatial activity of the reference image. After determining the spatial activity, the coded image is multiplied with the masked image so that the distortion is weighted according to the details present in the coded image. The high frequency coefficients of the masked coded and masked reference images are compared to determine the amount of blurriness. The block diagram for the blurriness meter is explained in figure below.

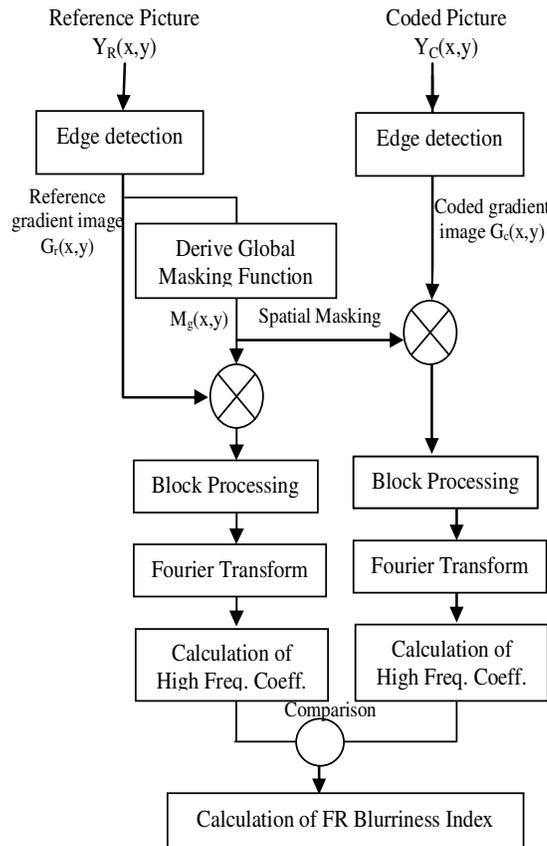


FIGURE 3: Block Diagram of blurriness detector

The blurriness estimation stage consists of 3 stages; Edge Detection; Spatial Masking; and analysis of high frequency coefficients.

First, the edge detection process is performed to determine the gradient of the reference and coded image. Then the edge detected version of the reference image is used to determine the spatial activity of the image for masking the blurriness distortion according to the local spatial activity of the reference picture in the form of masked image. Then the masked image is convolved with both, the edge detected reference image and edge detected coded image to weight the blurriness according to the spatial activity from the reference image. These steps are same as performed for the blockiness part.

Finally, the blurriness metric is calculated by comparing the high frequency coefficients of the masked reference and masked coded images in frequency domain. After calculation of the blockiness and blurriness indexes, the two distortions are combined together as explained in section below.

3.3 Combining Blockiness and Blurriness

The last stage of the blockiness and blurriness meter is to combine the two artifacts. As mentioned earlier, summation of blockiness and blurriness artifacts is not linear because both distortions have different visual impacts on the viewers. So first, they must be weighted accordingly and then added for a single quality metric.

To explain their combination phenomenon, 'bikes' image is compressed at different compression rates and the number of dominant blurred and blocky blocks (of 32×32 pixels) are calculated by comparing the reference and coded images.

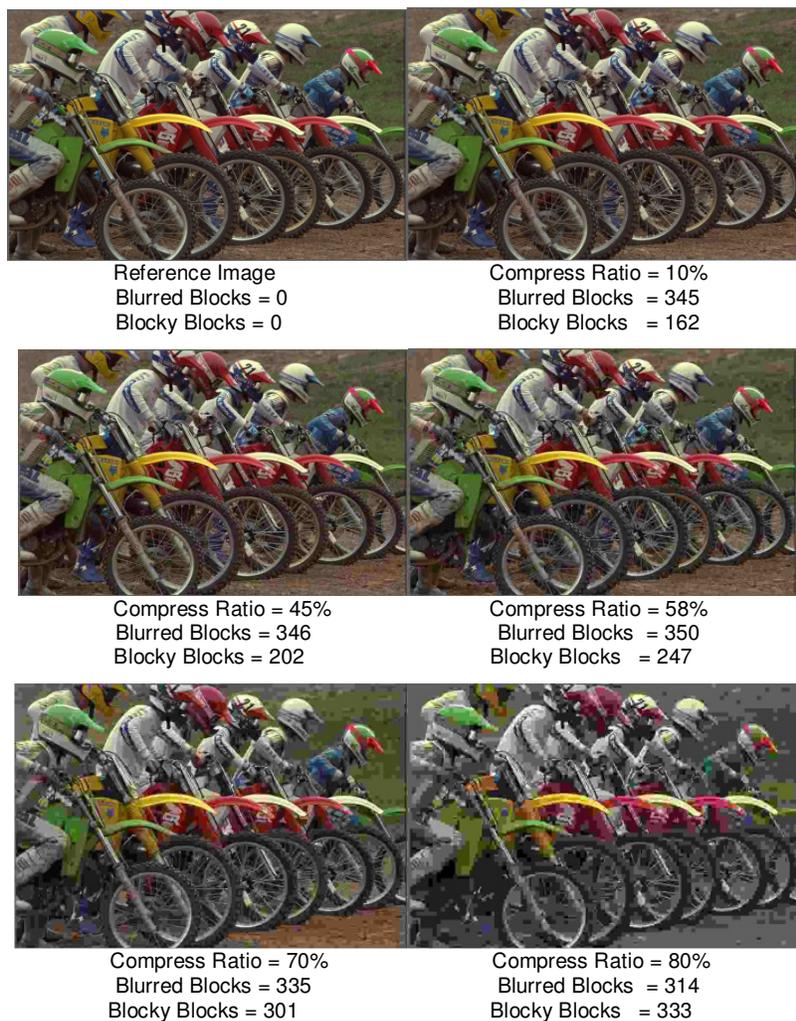


FIGURE 4: Bikes image compressed at different ratios

To understand the combination strategy of blockiness and blurriness artifacts we have to study their behavior and appearances in the images. As the compression ratio is increased, images tend to lose their higher frequency contents, due to their smaller energy they carry and appear blurry. When the compression becomes very severe, then the picture becomes blocky as can be seen in figure 4. This means, blockiness is an ultimate consequence of blurriness. Once the

blockiness starts appearing, it means the image has already gone through the blurriness artifact and the blurriness is saturated. By further compressing the image, blockiness artifacts starts appearing and it becomes dominant on blurriness artifact and user starts observing blockiness in image. Finally for the combination of two distortions, more weightage should be given to blurriness at low compressions and at higher compression rates to blockiness. The following graphs for blurriness and blockiness weighting functions are estimated based on tests on various images of the data base.

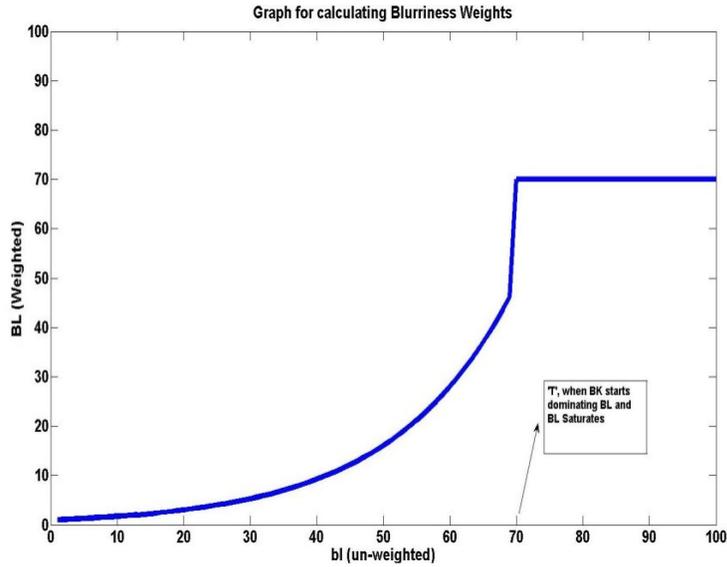


FIGURE 5: Blurriness weighting function graph

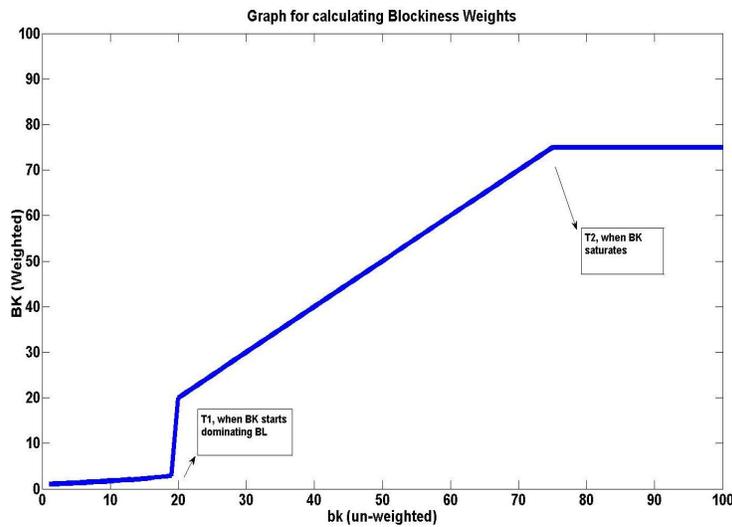


FIGURE 6: Blockiness weighting function graph

The above graphs highlight the combination of blockiness and blurriness artifacts. At low compression rates blurriness artifact appears so more weightage is given to blurriness at low compressions while the weightage of blockiness is more at high compression rates. The above distortion meter is tested on both blocky and blurred database from LIVE image database [11].

4. CONCLUSION

The main emphasis of the paper is to combine blockiness and blurriness distortions. As a fact that at low compressions blurriness comes first because of the loss of high frequency coefficients, after saturation of blurriness, blockiness artifact starts to appear as explained in figure 4. To combine these two artifacts, different weightage is given to both distortions. The complete algorithm was built in MATLAB software. The IQM is tested on blocky and blurred images from LIVE image database [11], and the results were compared with the users' Mean Opinion Score (MOS). The graph comparing MOS and IQM is shown below.

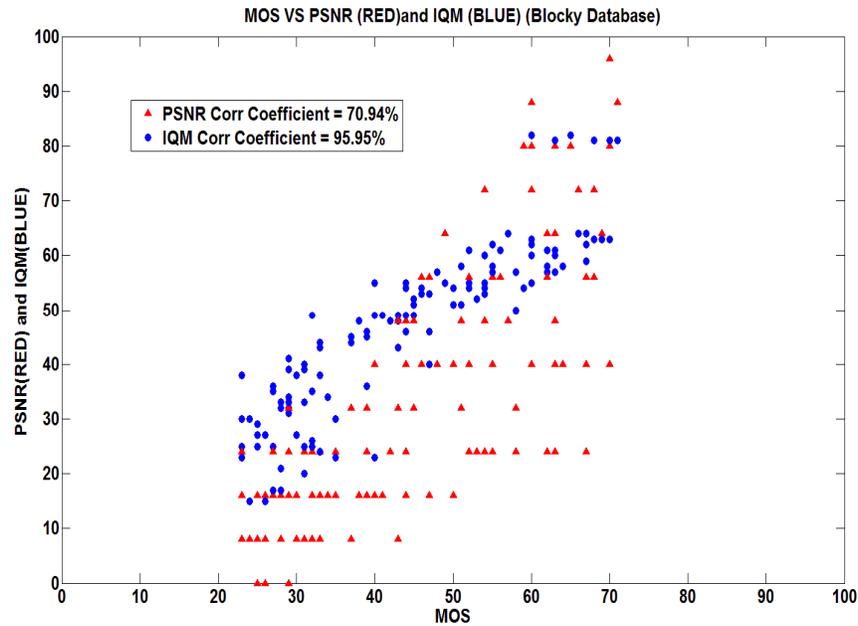


FIGURE 7: Comparison of the MOS and the objective quality for blocky image database from LIVE database

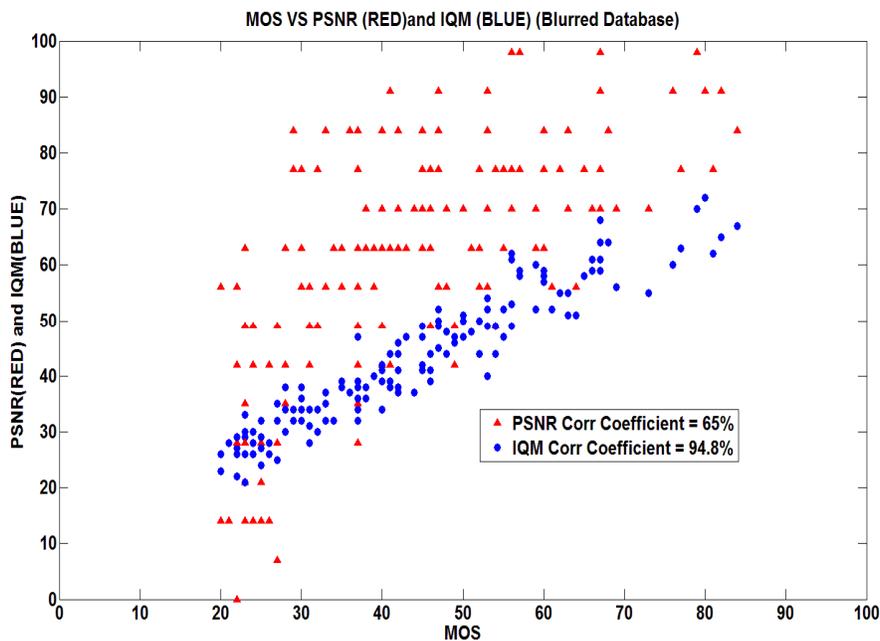


FIGURE 8: Comparison of MOS and the objective quality for blurred image database from LIVE database

Pearson's Correlation Coefficient (FR-Combined BK,BL) = 95.95 % (on blocky image database)

Pearson's Correlation Coefficient (FR-Combined BK,BL) = 94.8 % (on blurred image database)

5. REFERENCES

- [1]. Z. Wang and E. P. Simoncelli, "Reduced-Reference Image Quality Assessment Using A Wavelet-Domain Natural Image Statistic Model", Proceedings of SPIE Human Vision and Electronic Imaging X, vol. 5666, San Jose, CA, 2005, pp. 149–159.
- [2]. Z. Wang, A. C. Bovik, and B. L. Evans, "Blind measurement of blocking artefacts in images," in IEEE Int. Conf. Image Processing, September 2000, vol. 3, p. 981-984, IEEE.
- [3]. L. Meesters and J. B. Martens, "A single-ended blockiness measure for JPEG-coded images", Signal Process., Vol. 82, pp. 369-387, 2002.Z.
- [4]. M. Parvez Sazzad, Y. Kawayoke, and Y. Horita, "No-reference image quality assessment for jpeg2000 based on spatial features," Signal Processing: Image Communication, vol. 23, no. 4, pp. 257–268, April 2008.
- [5]. X. Zhu and P. Milanfar, "A no-reference sharpness metric sensitive to blur and noise", QoMEX, 2009. July 29–31, 2009, San Diego, California, U.S.A.
- [6]. M. G. Choi, J. H. Jung, and J. W. Jeon, "No-Reference Image Quality Assessment using Blur and Noise," in International Journal of Computer Science and Engineering 3:2 2009, pp. 76–80.
- [7]. R. Barland and A. Saadane, "A new reference free approach for the quality assessment of mpeg coded videos," in 7th Int. Conf. Advanced Concepts for Intelligent Vision Systems, Sep. 2005, vol. 3708, pp. 364–371.
- [8]. X. Feng and J. P. Allebach, "Measurement of ringing artifacts in jpeg images," in Proc. SPIE, Jan. 2006, vol. 6076, pp. 74–83.
- [9]. K. T. Tan, M. Ghanbari, "Frequency domain measurement of blockiness in MPEG-2 coded video," Proceedings of ICIP 2000, Vol. 3, pp.977 – 980.
- [10]. Fiorentini, A. and Zoli, M. T., "Detection of a Target Superimposed to a Step Pattern of Illumination. II. Effects of a Just-Perceptible Illumination Step", Atti. Fond. G. Ronchi, Vol. 22, pp. 207-217, 1967.
- [11]. Live website for subjective scores MOS.
<http://live.ece.utexas.edu/research/quality/>
- [12]. Z. Wang and A.C. Bovik, "A universal image quality index", IEEE Signal Processing Letters, 9(3):81–84, March 2002.
- [13]. Qadri, M.T. Tan, K.T. and Ghanbari, M., "Frequency domain blockiness measurement for image quality assessment ", IEEE International Conference on Computer Technology and Development (ICCTD), pp. 282 – 285, Cairo, 2-4 Nov. 2010.

Particle Swarm Optimization for Nano-Particles Extraction from Supporting Materials

M.A.Abdou

*Assistant Professor, Informatics Research Institute
City for Scientific Research and Technology Applications
P.O.Box 21934, Alexandria, Egypt*

m.abdou@pua.edu.eg

Abstract

Evolutionary computation for image processing is an encouraging research area. Transmission electron microscopy (TEM) images when used to characterize metallic and non-metallic nano-particles (size, morphology, structure, or composition), need such advanced image processing algorithms. This paper presents an efficient evolutionary computational method, particle swarm optimization (PSO), for automatic segmentation of nano-particles. A threshold-based segmentation technique is applied, where image entropy is attacked as a minimization problem to specify local and global thresholds. We are concerned with reducing wrong characterization of nano-particles due to concentration of liquid solutions or supporting material within the acquired image. The obtained results are compared with manual techniques and with previous researches in this area.

Keywords: Particle Swarm Optimization, TEM Images Scanning, Threshold Segmentation, Nano-Particles.

1. INTRODUCTION

Evolutionary computation algorithms (genetic-algorithms, genetic programming, or particle swarm optimizations) and data mining tools (neural network or decision tree) since introduced in image processing, have led to significant results [1-2-3]. In medical data classifications, neural networks [4-5] and linear programming models [6] obtained high classification accuracy rate, however their decision process was poor. Better results were achieved when using hybrid computation techniques, i.e. fuzzy-genetic, or neuro-fuzzy. In [7], a hybrid model was developed by introducing PSO for medical data classification.

Particle Swarm Optimization (PSO) is a relatively recent optimization method based on the idea of birds' swarming. PSO is similar to the Genetic Algorithm (GA) in the sense that they are both population-based search approaches. They both depend on information sharing among their population members to enhance their search processes. GA converges towards high quality solutions but within many and many iterations. PSO is easy to be implemented, have few parameters to be adjusted, and is more computationally efficient. This superior computational efficiency makes PSO more consistent in future image processing optimization problems [8].

From image processing point of view, TEM images characterization could be either manual or automatic. Existing techniques are facing different problems: illumination changes across the image, intensity variation of similar nano-particles due to diffraction contrast and/or liquids, and weak signal-to noise-ratio. TEM images contents' usually investigate binary or multi-class classifications [9]. Thresholding is an efficient image classifier tool, especially when real time processing is needed. Threshold selection can be either global or local. A global thresholding technique is one that segments the entire image with a single threshold value, whereas a local thresholding technique is one that partitions a given image into subimages and determines a threshold for each subimage [10]. Global techniques are further classified into point-dependent or region-dependent methods. When the threshold value is determined from the pixel gray tone independently from the gray tone of its neighborhood pixels, the thresholding method is point-dependent. On the other hand, a method is called region-dependent if the threshold value is determined from the local property within a neighborhood of a pixel.

1.1 Particle Swarm Optimization

PSO was originally introduced by Eberhart and Kennedy in 1995 [11]. In PSO algorithm, birds in a flock are our particles, where these particles are considered as simple agents through a problem space (food). PSO could be considered as an optimization technique. The algorithm starts by initializing a group of randomly distributed particles. These particles freely fly across the concerned space. During their flight, each initial particle should update its own velocity and position based on the experience of its own and the entire population. Instant particles' location in the multi-dimensional problem space represents one of the assumed solutions for the problem. When these particles move to new locations, another solution is generated. The velocity and direction of each particle will be recorded then altered in each generation. In any iteration, the particle's new location is computed by adding the particle's current velocity to its location. A fitness function is calculated for each solution and is to be minimized via further population of the swarms [8]. The PSO algorithm could better be summarized by means of a descriptive block diagram in the next section.

1.2 Nano-Particle Characterization Techniques in TEM Images

In [12], we have introduced a new fast Transmission Electron Microscopy (TEM) images clustering technique. In this research, analysis of particle sizes and shapes from two-dimensional TEM images were attacked. The hybrid method consisted of: an automatic segmentation and nano-particles counting. The automatic segmentation has assumed what we called "Automatic Threshold Generator" (ATG) towards a high efficient multiple- regions segmentation technique. ATG generated a vector of bi-threshold values used for electronic microscopic input image segmentation. This ATG gave good results when compared with existing algorithms [13]. However, TEM images where concentration of liquid solutions or supporting material affects image intensities failed to be counted via ATG correctly. Results were not comparable to manual results in these TEM images. This could be referred to wrong classification of supporting materials as nano particles.

2. The Proposed Nano-particle Segmentation Algorithm

In this paper we are attacking threshold-segmentation problem as a minimization problem, where both local and global minima are to be detected via an evolutionary computational method (PSO). The objective is to search for a better classifier that differentiates between nano-particles and their supporting materials.

2.1 PSO Algorithm

Figure (1) presents a summary of the PSO steps. Equation (1) refers to a number 'rand' which is a generated positive random number less than unity. As we are dealing with computer based simulation, the variable Δt is generally referred to the iterations span that could be considered unity. Thus, X_0^i is limited between a minimum and a maximum value. The second step is the particles velocity update through the iterative procedure. In equation (2), the particles velocities are updated (V_{k+1}^i) looking towards the initial velocity (V_k^i) and influences emerging from: the particle itself (p_i) and the swarm (p_k^g); where p_i denotes the best position required from the fitness function for the particle itself and p_k^g is the best position obtained globally for any particle within the whole swarm towards the desired minimum. The values of w , c_1 , and c_2 are all knowledge based constants that take values from experience knowledge. Equation (3) explains the position updates. Equation (4) deals with the memory update for the fittest position and velocity. Finally, equation (5) concerns with the stopping criteria. The algorithm could stop in one of two conditions: either the error obtained is less than the required error, or the algorithm undergoes successive iterations without any improvement in the error.

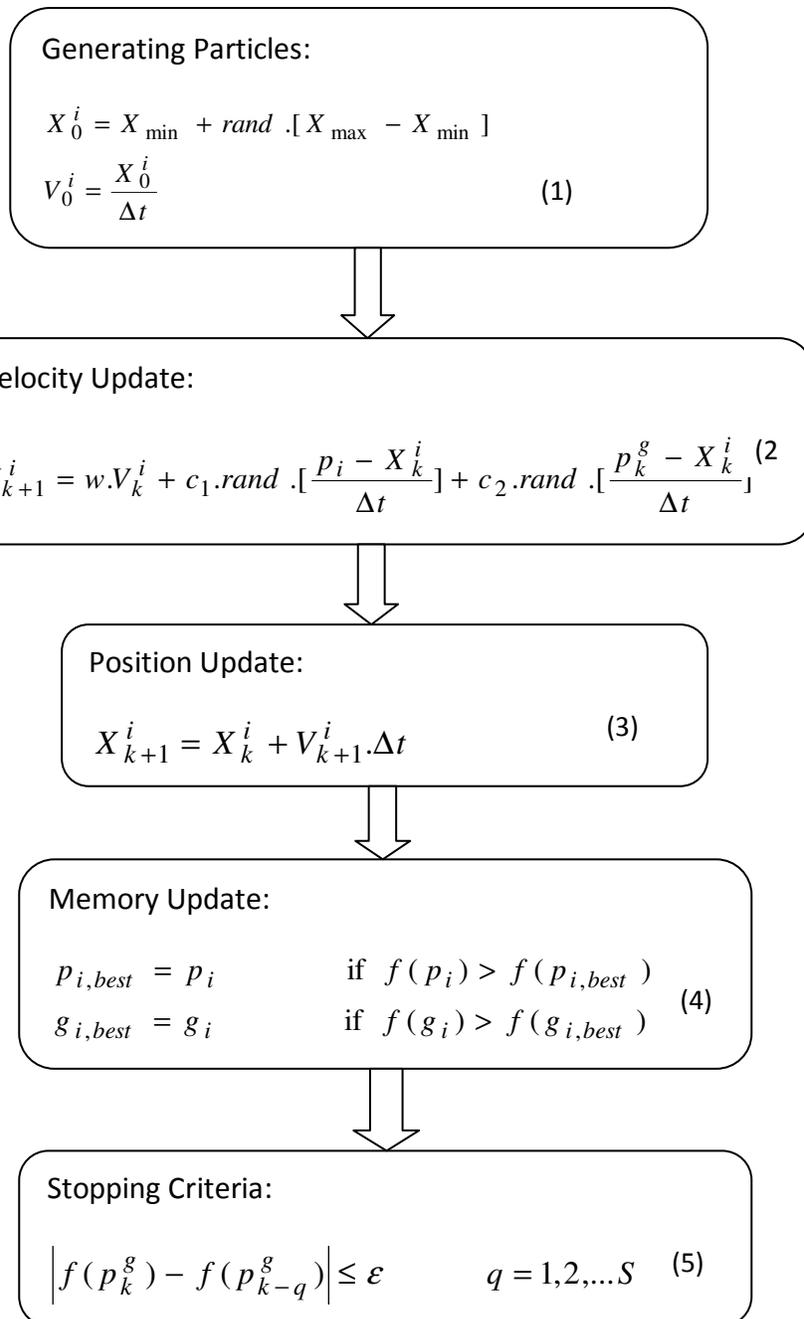
2.2 Entropy- Based Segmentation

The motivation of application of the maximum entropy method to solve threshold selection has been started since 1989 [14]. The maximum entropy principle states that, for a given amount of information, the probability distribution which best describes our knowledge is the one that maximizes the Shannon entropy subjected to a set of constraints [15]. The Entropy mathematical model could be summarized as figure (2). This model assumes first the number of classes required for the nano-particles image. We have selected two thresholds in order to compare the obtained results with the ATG previously presented in

[12]. Selecting two threshold means that we have three classes: $[0, t_1]$ where we will define an entropy H_0 ; from $[t_1, t_2]$ where entropy H_1 is defined, and finally $[t_2, L]$ where H_2 is calculated. 'L' represents the maximum grayscale obtained during histogram analysis.

2.3 Thresholds Generation

The previously discussed PSO algorithm is included within a segmentation procedure, as shown in figure (3). This algorithm starts with a histogram generator, a PSO implemented algorithm where the target function is the entropy and the output is the threshold vector. Images are then segmented to separate the nano-particles from the TEM image. Swarming, the PSO technique generates a vector of bi-threshold values used for electronic microscopic input image segmentation.



The advantage of the PSO when compared to other existing techniques concerns with the dimensionality broadband of the PSO. This means that PSO could generate a multi-threshold values vector with n -dimensions according to the type of images to be segmented. The vector length is related to the entropy equation applied which in turns is related to the degree of the image classifier.

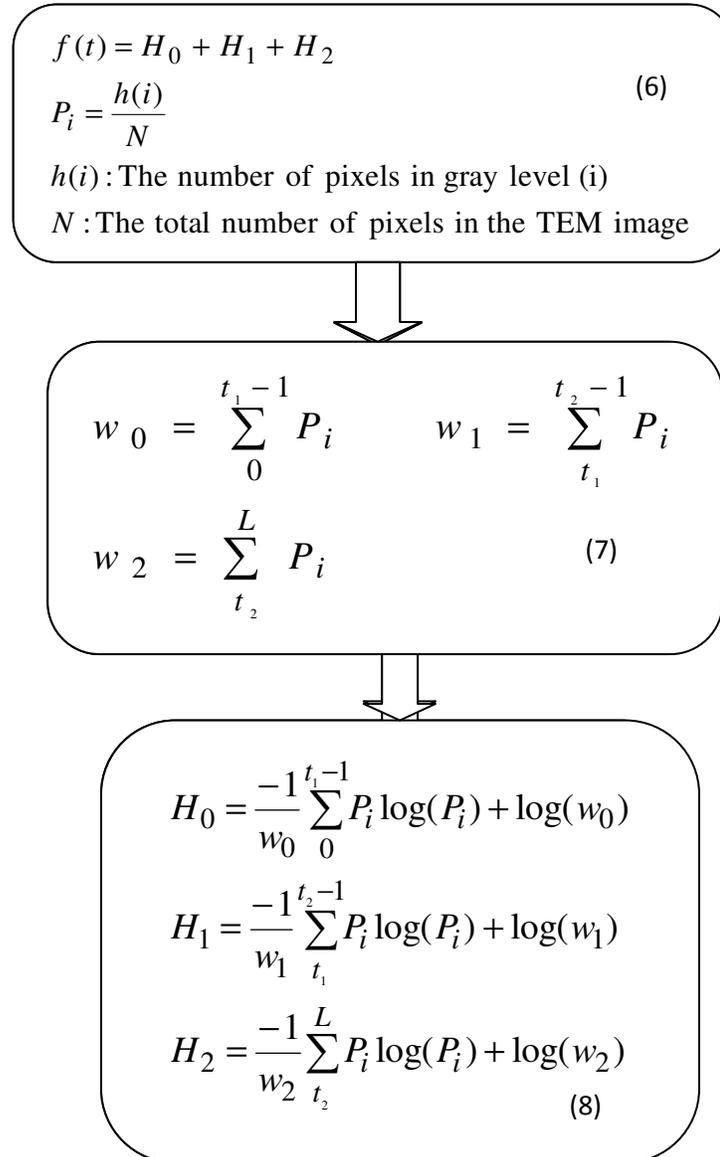


FIGURE 2: Entropy Calculations of TEM Image.

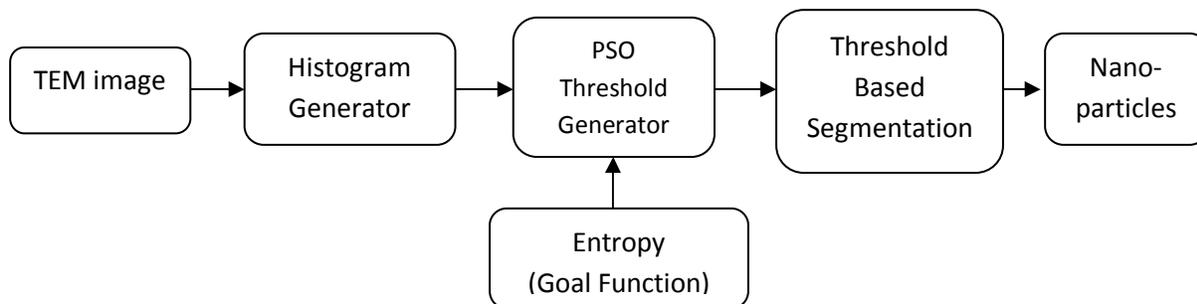


FIGURE 3: The Segmentation Technique: The PSO uses entropy as a goal function.

3. Results and Discussion

The introduced algorithm is applied to a wide TEM data set. Comparison starts with cases described in [12] to evaluate the improvement gained. Table (1) shows a comparison between this PSO based segmentation and the previous presented ATG method [12] for four different selected cases: nano1- nano2- nano3- nano4. Table (2) shows the ratio of nano-particles characterized by means of two different algorithms. The improvement in the PSO method is remarkable in cases: nano3 and nano4. Figures (4 and 5) show a comparison between two easy segmented cases. PSO gives results approximately equal to those obtained in [12].

Case	ATG [12]		Proposed PSO	
Nano1	T1 = 0	T2 = 73.5	T1 = 0.0441	T2 = 70.1350
Nano2	T1 = 0	T2 = 83.5	T1 = 1.0643	T2 = 70.449
Nano3	T1 = 0	T2 = 127	T1 = 1.1497	T2 = 50.1398
Nano4	T1 = 0	T2 = 64	T1 = 0.1290	T2 = 30.1241

TABLE 1: Comparison between the two entropy-based segmentation techniques.

Case	Ratio of nano-particles extracted		Improvement
	ATG [12]	Proposed PSO	
Nano1	0.0401	0.0356	0.4532
Nano2	0.1015	0.0771	2.4475%
Nano3	0.2527	0.0358	21.6934%
Nano4	0.2843	0.0133	27.1027%

TABLE 2: Nano-particles extraction effectiveness of the proposed method.

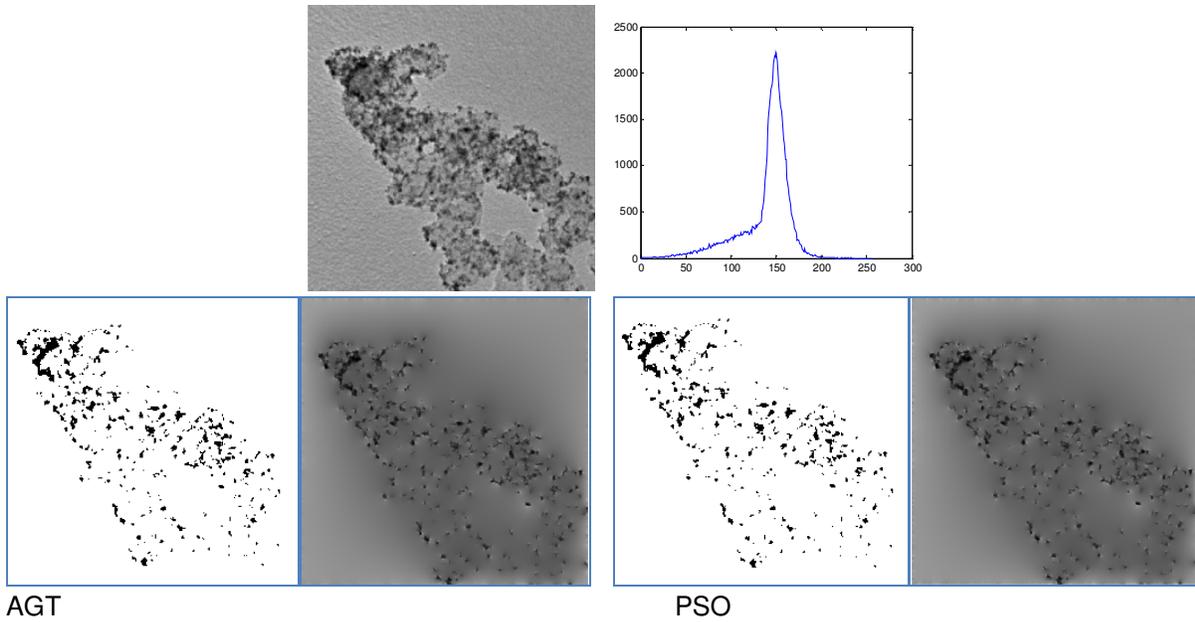


FIGURE 4: Nano1 case: Original image- Histogram- AGT results- PSO results

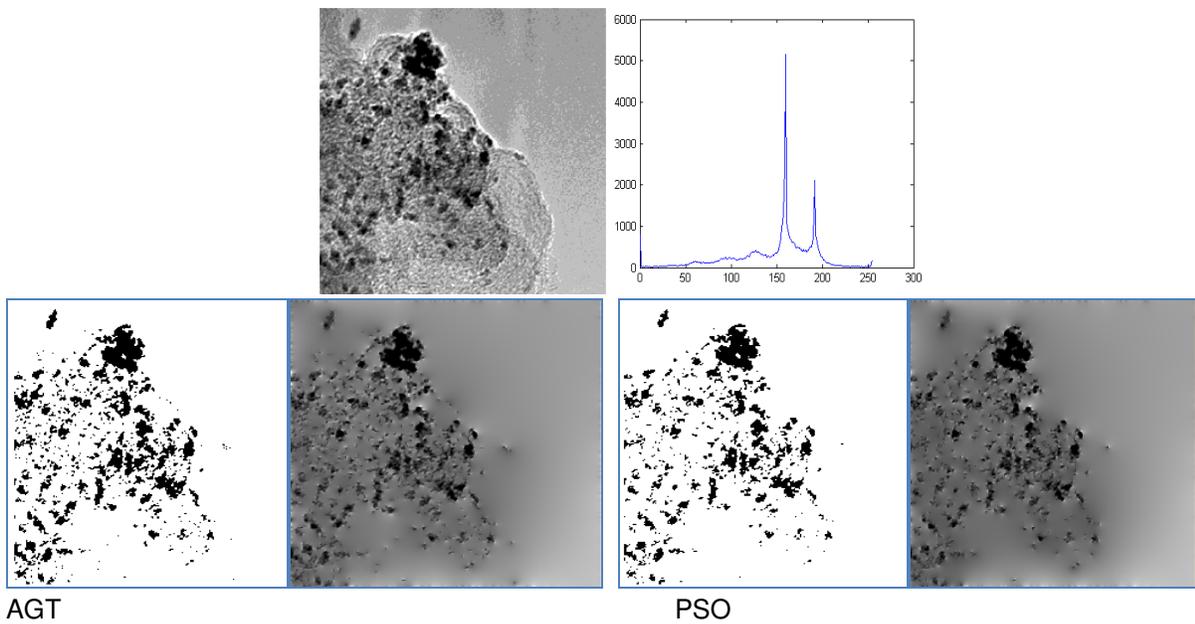


FIGURE 5: Nano2 case: Original image- Histogram- AGT results- PSO results.

Figures (6 and 7) show how the proposed PSO segmentation succeeded in segmenting smaller regions corresponding to nano-particles areas. False areas are characterized as nano-particles due to fluids within the specimen sample. In figure (8), the false areas characterized as nano particles using the ATG [12] are shown. All these areas are not the required targeted particles. The histogram presented in figure (9) is a comparison of one manual and two automatic counting methods. The PSO presented in this work is close to the manual method especially in images that are difficult to be segmented.

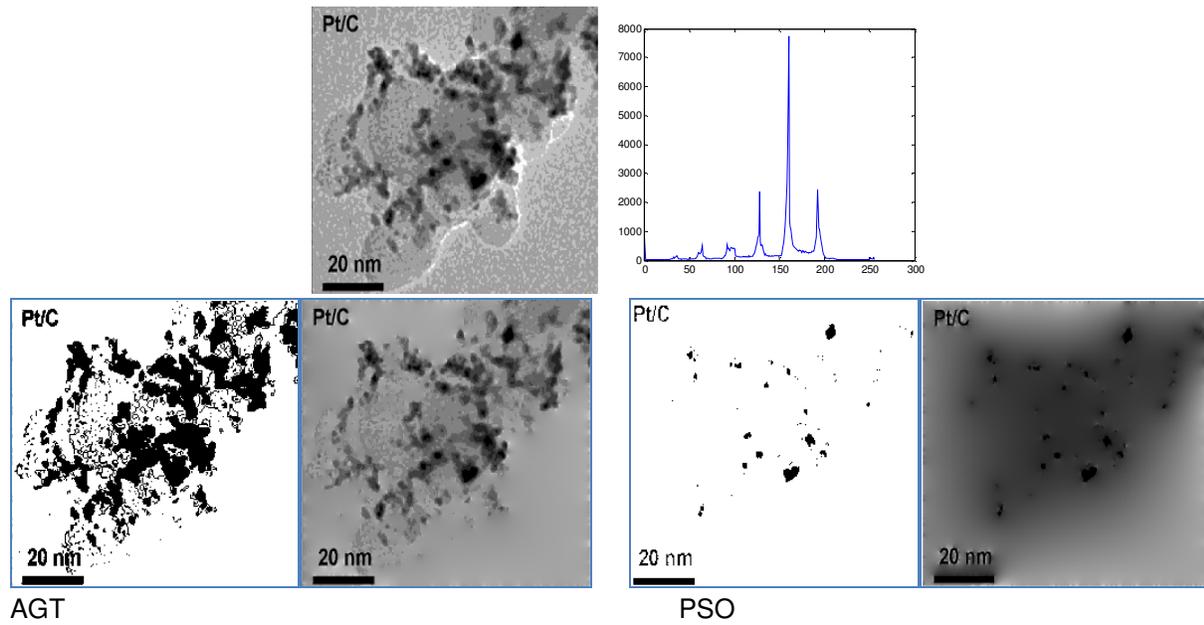


FIGURE 6: Nano3 case: Original image- Histogram- AGT results- PSO results

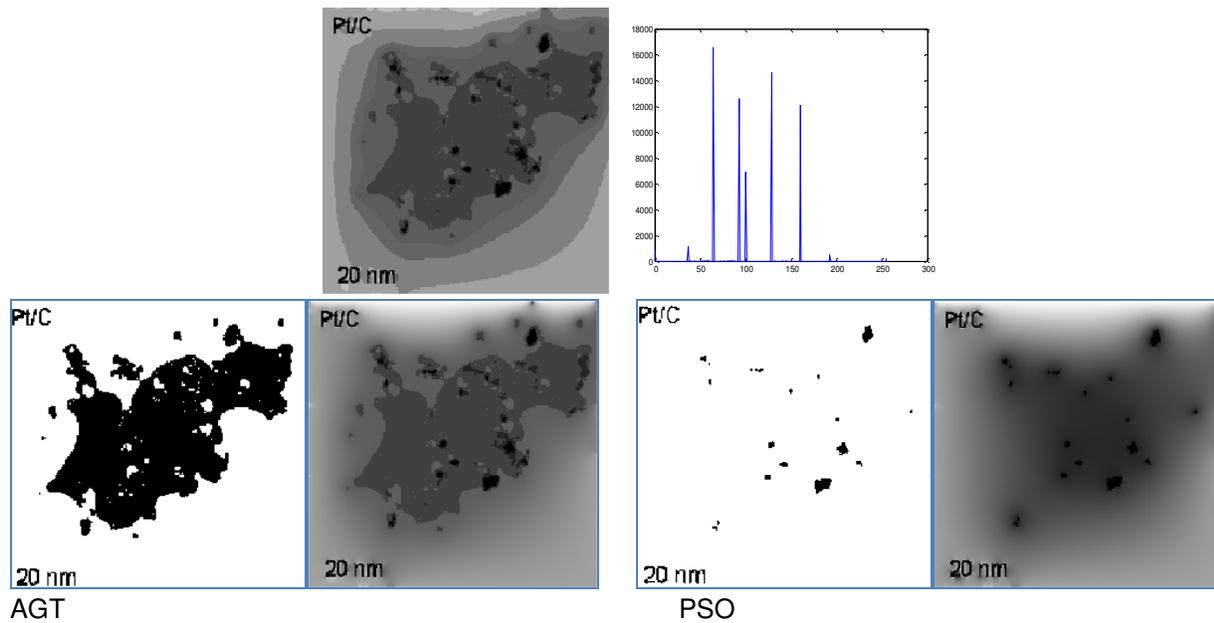


FIGURE 7: Nano4 case: Original image- Histogram- AGT results- PSO results

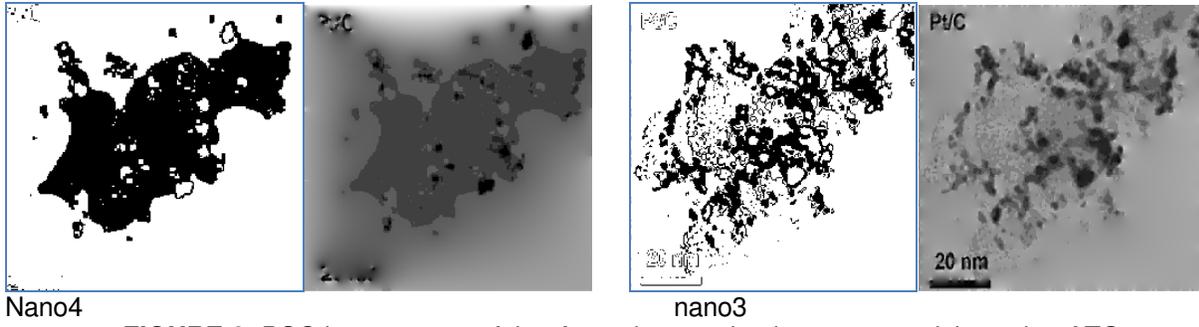


FIGURE 8: PSO improvement: false Area characterized as nano-particles using ATG.

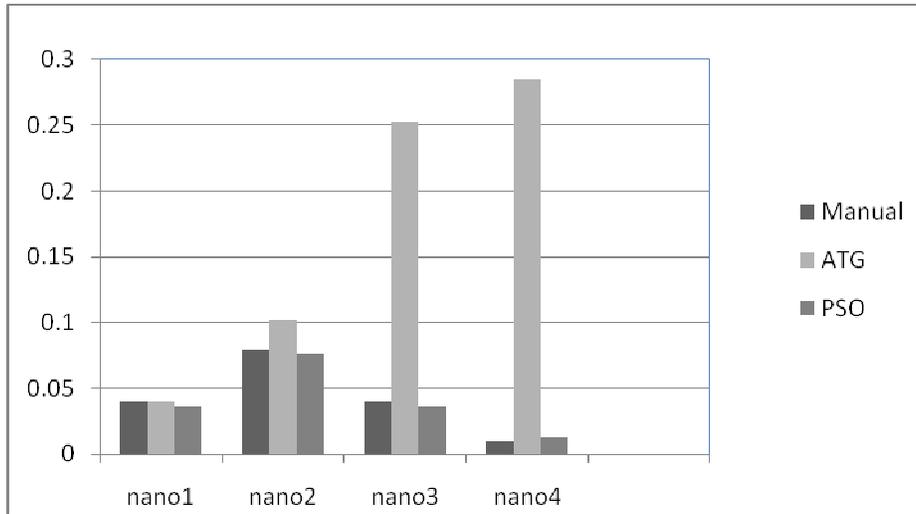


FIGURE 9: Nano-particles counting: Manual- ATG [12]- Proposed method.

CONCLUSION

The PSO- automatic segmentation presented has specified a threshold vector that clusters the nano-particles existing within the TEM image, where image entropy measure is the goal function. Obtained results show that the proposed segmentation method reduces wrong characterization of nano-particles in images where concentration of liquid solutions or supporting material affects image intensities. The PSO segmentation surpasses the ATG, presented in [12], by about 27%. Furthermore, the counted particles although comparable to manual results, the PSO presented technique has shown high computational efficiency which serves for real time characterization.

REFERENCES

- [1] C.A. Pena-Reyes, M. Sipper. "Evolving fuzzy rules for breast cancer diagnosis". In Proceedings of 1998 International Symposium on Nonlinear Theory and Applications (NOLTA'98), 2: 369–372, 1998.
- [2] C.A. Pena-Reyes, M. Sipper. "A fuzzy-genetic approach to breast cancer diagnosis". Artificial Intelligence in Medicine, 17: 131–155, 1999.
- [3] X. Chang, J.H. Lilly. "Evolutionary design of a fuzzy classifier from data". IEEE Transactions on Systems, Man, and Cybernetics, 34: 1894–1906, 2004.
- [4] R. Setiono. "Extracting rules from pruned neural networks for breast cancer diagnosis". Artificial Intelligence in Medicine, 8: 37–51, 1996.
- [5] R. Setiono, H. Liu. "Symbolic representation of neural networks". Computer, 29: 71–77, 1996.
- [6] K.P. Bennett, O.L. Mangasarian. "Neural network training via linear programming". Advances in Optimization and Parallel Computing, 56–67, 1992.
- [7] Pei-ChannChang, Jyun-JieLin, Chen-HaoLiu. Computer methods and Programs in Biomedicine, Vol.(not yet): 2011.
- [8] L.Zhang, T.Mei, Y.Liu, D.Tao, H.Zhou. "Visual search reranking via adaptive particle swarm optimization". Pattern Recognition, 44: 1811- 1820, 2001.
- [9] C.C. Bojarczuk, H.S. Lopes, A.A. Freitas. "Genetic programming for knowledge discovery in chest-pain diagnosis: exploring a promising data mining approach". IEEE Engineering in Medicine and Biology Magazine 19: 38–44, 2000.
- [10] Du Feng, Shi Wenkang, Chen Liangzhou, Deng Yong, Zhu Zhenfu. "Infrared image segmentation with 2-D maximum entropy method based on particle swarm optimization (PSO)". Pattern Recognition Letters, 26: 597–603, 2005.
- [11] J Kennedy, R. Eberhart. "Particle Swarm Optimization". 1942–1948, (1995).
- [12] M.A. Abdou, Bayumy B.A. Youssef, W.M. Sheta. "Nano-particle Characterization Using a Fast Hybrid Clustering Technique for TEM Images". (IJCSIS) International Journal of Computer Science and Information Security, 8 (9): 101-110, 2010.
- [13] T.Chem, Tubitak, H.Woehrle, E.Hutchison, S.Ozkar, G.Finke. "Analysis of Nanoparticle Transmission Electron Microscopy Data Using a Public- Domain Image-Processing Program". Image Metrology, 30: 1-13, 2006.
- [14] N.R.Pal, S.K.Pal, Object-background segmentation using new definitions of entropy, Proc. Inst. Elec. Eng. 136, pp.284-295, 1989.

- [15] G.J.Klir, T.Folger. "Fuzzy Sets Uncertainty and Information". Printice Hall, Englewood Cliffs, NJ, (1988).

INSTRUCTIONS TO CONTRIBUTORS

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

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

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

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

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

LIST OF TOPICS

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

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

- Storage and transmission

- Video coding algorithms and technologies for ATM/p

CALL FOR PAPERS

Volume: 6 - **Issue:** 1 - February 2012

i. Paper Submission: November 30, 2011

ii. Author Notification: January 01, 2012

iii. Issue Publication: January / February 2012

CONTACT INFORMATION

Computer Science Journals Sdn Bhd

B-5-8 Plaza Mont Kiara, Mont Kiara

50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6207 1607

006 03 2782 6991

Fax: 006 03 6207 1697

Email: cscpress@cscjournals.org

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

PHONE: 006 03 6207 1607
006 03 2782 6991

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