

# International Journal of Image Processing (IJIP)



ISSN : 1985-2304



VOLUME 4, ISSUE 5

PUBLICATION FREQUENCY: 6 ISSUES PER YEAR

# **International Journal of Image Processing (IJIP)**

**Volume 4, Issue 5, 2010**

**Edited By**  
**Computer Science Journals**  
[www.cscjournals.org](http://www.cscjournals.org)

**Editor in Chief Professor Hu, Yu-Chen**

# **International Journal of Image Processing (IJIP)**

Book: 2010 Volume 4, Issue 5

Publishing Date: 30-10-2010

Proceedings

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. Violations are liable to prosecution under the copyright law.

IJIPJournal 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**

## **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 fifth 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.

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)*

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

**Professor. Davide La Torre**  
*University of Milan (Italy)*

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

**Dr. Huiyu Zhou**  
*Queen's University Belfast (United Kingdom)*

**Professor Yen-Wei Chen**  
*Ritsumeikan University (Japan)*

## Editorial Board Members (EBMs)

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

**Professor. Herb Kunze**  
*University of Guelph (Canada)*

**Professor Karray Fakhreddine**  
*University of Waterloo (United States of America)*

**Assistant Professor. Yufang Tracy Bao**  
*Fayetteville State University (North Carolina)*

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

**Dr. Ghassan Adnan Hamid Al-Kindi**  
*Sohar University (Oman)*

**Dr. Cho Siu Yeung David**  
*Nanyang Technological University (Singapore)*

**Dr. E. Sreenivasa Reddy**  
*Vasireddy Venkatadri Institute of Technology (India)*

**Dr. Khalid Mohamed Hosny**  
*Zagazig University (Egypt)*

**Dr. Gerald Schaefer**  
*Loughborough University (United Kingdom)*

**Dr. Chin-Feng Lee**  
*Chaoyang University of Technology (Taiwan)*

**Associate Professor. Wang, Xiao-Nian**  
*Tong Ji University (China)*

**Professor. Yongping Zhang**  
*Ningbo University of Technology (China)*

**Professor Santhosh.P.Mathew**  
*Mahatma Gandhi University (India)*

# Table of Content

Volume 4, Issue 5, December 2010

## Pages

- 457-467 HSV Brightness Factor Matching for Gesture Recognition System  
**Mokhtar M. Hasan, Pramod K. Mishra**
- 468 -475 A Novel Secret Sharing Technique Using QR Code  
**Jun-Chou Chuang, Yu-Chen Hu, Hsien-Ju Ko**
- 476-483 Automated Monitoring System for Fall Detection in the Elderly  
**Shadi Khawandi, Bassam Daya, Pierre Chauvet**
- 484-508 Image Fusion and Image Quality Assessment of Fused Images  
**MANJUSHA, UDHAV**
- 509-517 A novel Image Retrieval System using an effective region based shape representation technique  
**Santhosh.P.Mathew, Philip Samuel**

# HSV Brightness Factor Matching for Gesture Recognition System

**Mokhtar M.Hasan**

*Department of Computer Science  
Faculty of Science  
BHU-Varanasi  
Uttar Pradesh-221005, India*

mmwaeli@gmail.com

**Pramod K. Mishra**

*Department of Computer Science  
Faculty of Science  
BHU-Varanasi  
Uttar Pradesh-221005, India*

pkmisra@gmail.com

---

## Abstract

The main and primary objective of the gesture recognition research is to establish a system which can identify specific human gestures and utilize these identified gestures to be carried out by the human made machines, In this paper, we are going to introduce a new method for gesture recognition that based on the local brightness of each block of the gesture image, the input gesture image is divided into 25x25 blocks each of 5x5 block size, and we have calculated the local brightness of each divided block after applying colored segmentation operation using HSV (Hue, Saturation and Value) color model, so, each gesture produces 25x25 feature values called feature vector, our experimental shows that more that 65 % of these features are zero value which leads to minimum storage space, the recognition rate achieved is 91 % using 36 training gestures and 24 different testing gestures. We build a gesture recognition system that can communicate with the machine in natural way without any mechanical devices and without using the normal input devices which are the keyboard and mouse and the mathematical equations is the only translator that existed between the human and human-made machines, we have focused in this study on the hand gesture since hand can represents more meaning than other human-bodily organs.

**Keyword:** Brightness Calculation, HSV color model, Gesture Recognition, Template Matching, Image Segmentation, Laplacian Edge Detection.

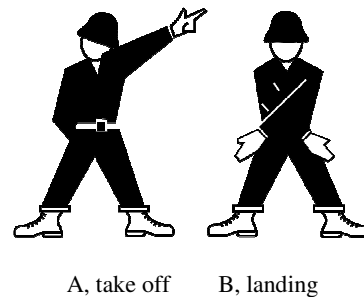
---

## 1. INTRODUCTION

In all of the simulation processes, we tried to simulate the human abilities, in gesture recognition system, the remarkable ability of the human vision is the gesture recognition, it is noticeable mainly in deaf people when they communicating with each other via sign language and with hearing people as well. In this paper we tried to simulate this ability but this time will be between the human and human-made machines.

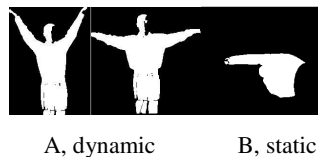


A gesture pose is a mode of communication between people that depends on the bodily movement especially the hand motion and pose; this form of communication is established along with spoken words in order to create a comprehensive statement to be carry out be the hearer. Most people use gestures language represented by bodily movement in addition to spoken words when they communicate between each other [1]; Figure (1) shows a gesture example for helicopter signaler.



**FIGURE 1:** Helicopter Signaller for Marshaling Operations [2].

The normal communication between the people is the speaking which needs the sound to convey the meaning, while the later kind needs the space to convey the meaning [3]. The coarse classification of gestures is two; static and dynamic, the static gesture is a specific hand pose formed by a single image. The dynamic gesture is a moving gesture formed by a sequence of images [3] as in Figure (2),



**FIGURE 2:** A and B Represent the Dynamic and Static Gesture Respectively.

The application of gesture system on interactive applications produces many challenges. The first and important challenge is the response time which should be fast [4]. There should be no noticeable time between user gesture movement and computer replies [4]. The designed computer vision algorithms should be reliable and work for different ethnic people [4] especially when the color of human is changed comparing with white and black people. One more challenge which is the cost challenge, the gesture system needs special hardware such as the camera and sensors as necessarily, those special hardware will be the replacement of the existing hardware devises which may considered as low cost [4] such as the keyboard and mouse, but the gesture system with these new devices will be more worthwhile for wire-less communication.

This paper applied a new gesture recognition method for identifying the gestures for the computer or for the telerobotic in order to understand and carry on the human teleoperations, we have applied this novel method by windowing the image in order to recognize the input gesture and discover the meaning for that gesture. We have applied the proposed method using six gestures database, each of ten samples, so the total is sixty gestures database used for gesture recognition; we used the hand gesture rather than the face because the hand is the most flexible part of the body and can shows different meaning.

## 2. RELATED WORK

Roberto and Tomaso [5] applied a face recognition template matching; a very simple and direct recognition technique based on the use of whole image as grey-level templates. The most direct of the matching procedures is correlation. First, the image is normalized to obtain unified locations for the mouth, eye and nose. The authors in [5] applied her technique by creating a database entry for each person contains the frontal view of that person, along with four masks which are the eyes, nose, mouth, and face (the region from eyebrows downwards as decided by the authors in [5]). All these four masks are relatively to the position of normalized eye position in whole of their database. The recognition measure applied by them is the Euclidian Distance; by matching the new presented face gesture with all the database gestures, and the database gesture corresponds to the maximum matching score is the recognized gesture, they had used samples taken from 47 persons with 4 gestures each.

Freeman and Roth [6] applied hand gesture recognition using orientation histogram, They had applied some transformation T to the image data in order to create the feature vector that will be used for recognition purpose and represents that specific gesture. To classify the gesture, they compare the feature vector with the feature vectors from a previously generated training set. The transformation T can be described as a polar representation for the histogram of the local orientations of the input gesture, they use the gradient direction to calculate this orientation, the histogram of the directions is then sketched using polar plot which represents the final features of the input gesture and the other features will treated the same, they had used samples taken from one person with 5-15 gestures.

K. Symeonidis in [4] applied gesture recognition method using neural network, he has used 8 different hand gestures each of 3 samples, those were for training purpose, he did not use an exact number of gestures for testing purpose since some gestures tolerate more samples than others, the features than he has used were the 19 elements degree numbers that been converted from the polar representation of the orientation histogram of the input gesture, then he presents these features for training the neural network after some preprocessing that casted the original features into later number of features.

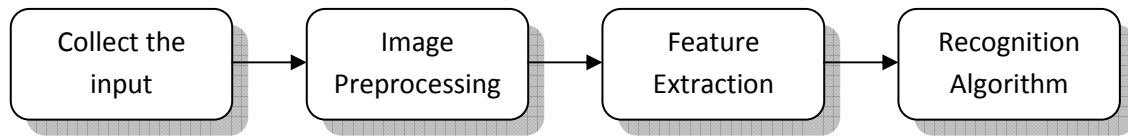
Xingyan [7] applied gesture recognition using fuzzy C-means algorithm, He has used image processing methods to transform the raw image into the feature vector. The feature vector is created by applying segmentation using HSV color model and then he has reduced the noise, the feature vector of the image is thirteen parameters long. The first feature is the aspect ratio of the hand's bounding box as decided by the authors in [7]. The last 12 features are values representing coarse parameters of the image, where each grid cell is the mean gray level value in the 3 by 4 block division of image. Each of the 12 values calculated by the mean value of 3 by 4 partitions, which represent the mean of the brightness value. The classification phase is applied using a recognition algorithms based on the Fuzzy C-Means (FCM) algorithm, he had used samples taken from 6 persons with 6 gestures each and achieves a time of 2-4 seconds with 86 % recognition rate.

## 3. OVERALL APPROACH

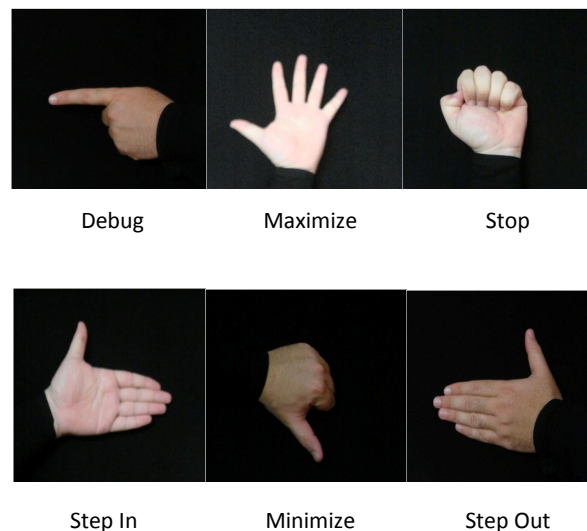
Our system is composed of four main segments in order to recognize the input gesture, these stages are summarized in Figure (3), and the details of these stages are as follows:

### 3.1 Collect the Input

In this phase, the input gestures are predisposed and the gesture poses are decided along with their meaning, and the database is creating which contains the different gestures poses with many samples per gesture, the number of the samples are limited by the speed and accuracy of the system, as the samples increased, the system speed decreased and the accuracy increased, and vice versa. We have chosen our database as in Figure (4) along with their meaning.



**FIGURE 3:** Overview of Gesture Recognition System.



**FIGURE 4:** System's Database Vocabulary.

As seen by Figure (4), the database contains six gesture poses, which represent the target recognition decision which may belong to any of these six gestures if the new presented gesture is recognized, else, the system will announce that the new presented gesture is not-known, these six gestures that implied in the database have many samples for each gesture which are six samples for each, as the number of samples increases the system accuracy increases and the testing time increases which effects badly on the overall speed and performance of the system.

### 3.2 Image Preprocessing

In this phase, we have predisposed the input gestures for both training and testing purposes; this phase is divided down into the following sub phases.

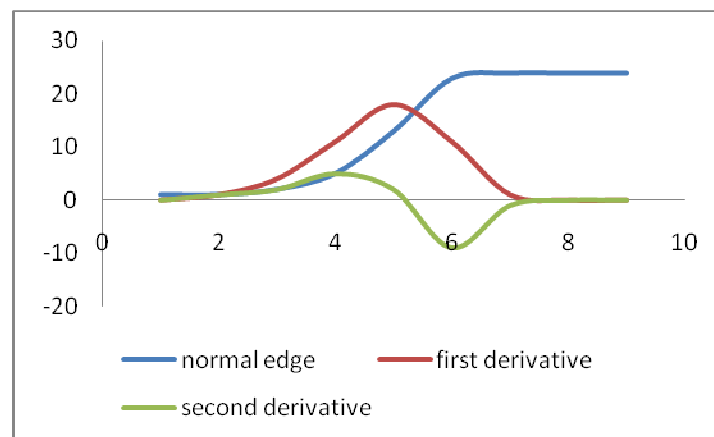
#### 3.2.1 Segmentation

In this phase, we have applied the segmentation operation to segment the hand area in the input gesture and isolate it from the background, all of the gesture systems depend on the perfect segmentation of the hand gesture region, there are two main methods for segmentation: first method is by using HSV model; which deals with the color pigment of the human skin, the different ethnic groups have a significant property which is the different in the skin color which is represented by the pigment concentration difference which affect the saturation of the skin [7]. The color of the skin, on the other hand, is roughly invariant across ethnic groups [7]. By deciding a range for each of H, S and V parameters, a good segmentation method can be achieved; Xingyan Li [8] have decided certain values for H and S parameters only. The other method used for segmentation operation is by using clustering algorithms or thresholding technique, these algorithms are suitable mainly for homogeneous or uniform background that has no cluttered

objects and the hand is the prominent object, we have applied HSV mode of segmentation for splitting the hand area from the image gesture with some threshold value, after image segmentation; we have to normalize the segmented image gesture in order to obtain gestures database that are invariant against position, scale, and rotation, this will speed up the recognition operation and reduce the number of gesture samples stored in the database.

### 3.2.2 Edge Detection

The areas in images with high contrast that has an intensity jump between the connected pixels are called the edges, the non-edge areas do not has any intensity difference, noise also has some intensity difference, so, edge detection algorithms preserve the important information in the images that represents the edges and reduces the useless information which is classified as non-edges pixels, there are two types of edge methods, the gradient (first derivative) and the laplacian (second derivative) as Figure (5), the gradient seeking for the maximum and minimum values and by using of thresholding techniques it classifies the edges and this methods is used when the intensity is high and is changed rapidly which provide a significant contrast, but when the gray level changes slightly from dark to light or vice versa the second derivative is used; the second derivative for maximum value is zero value which is the principle of the laplacian edge detector, this methods uses zero crossing to find the edges and it is suffers from false edges produced because of noise so it need special treatment like blurring the image before applying laplacian edge detector, we used this laplacian to find the edges.



**FIGURE 5:** Edge Methods Categories.

As seen by Figure (50), the maximum value hereinabove is the edge in case of first derivative methods; this edge can be located by using of thresholding technique which produces different edge location depending on the threshold value and the edge thinness depends on the correct selection of the threshold as well, but in second derivative; the intersection with x-axis is the edge produces a unique and non-duplicated edge.

### 3.2.3 Normalization

In this phase of image preprocessing, the gesture is trimmed to get rid of the unnecessarily area that surrounding the gesture area, this is done by removing this useless area from four directions.

### 3.3 Feature Extraction

After the preparing of the image and segmenting of the hand gesture, a black-white image is created and represented the hand pose inset, the feature extraction phase will start, the overall feature vector size is 625 elements which are the brightness values of each block in the gesture, these features are stored in the database and the same algorithm is applied on the 36 gesture database, the actual feature vector size that stored in the database is more smaller than the calculated feature vector since the existence of zero values, this fact will be discussed hereafter.

### 3.4 Recognition Algorithm

When a new input presents to the system for classification, the input gesture is treated as prescribed sequence and extracting the features of this input gesture, the feature vector of the new presented input is compared against the stored database feature vectors using our new algorithm for matching purpose, each two brightness values are considered equal in case of both have no brightness value which means the block which is 5x5 (25 pixels) is black and this lead us to matching between two black areas, and in case of the existence of the brightness value, i.e.  $>0$ , we apply some threshold for considering the two blocks are equal, we use threshold of one to obtain maximum flexibility of the system in case of changing the inset of the same gesture, this matching algorithm yields good recognition rate comparing with other prescribed methods because the recognition rate depends on two decisions, the black matching and the brightness matching.

## 4. EXPERIMENTAL RESULTS

As we explained before, we trained the system with six different gestures each of six samples, these gesture images have undergone in a serial operations in order to extract the features.

First step is the segmentation operation, this operation is required for splitting the hand region, the segmentation is applied using HSV color space, the input gesture is RGB color space and converted to HSV color space using the following equations:

Let  $p(x, y)$  represents the input pixel with R, G, and B components, and let  $p'(x, y)$  with H, S, and B components, the following steps for converting  $p(x, y)$  to  $p'(x, y)$  [12]:

Step1: calculate M from,  $M = \max(R, G, B)$

Step2: calculate m from,  $m = \min(R, G, B)$

Step3: calculate r from,  $r = (M-R)/(M-m)$

Step4: calculate r from,  $g = (M-G)/(M-m)$

Step5: calculate r from,  $b = (M-B)/(M-m)$

Step6: calculate V output from,  $V = \max(R, G, B)$

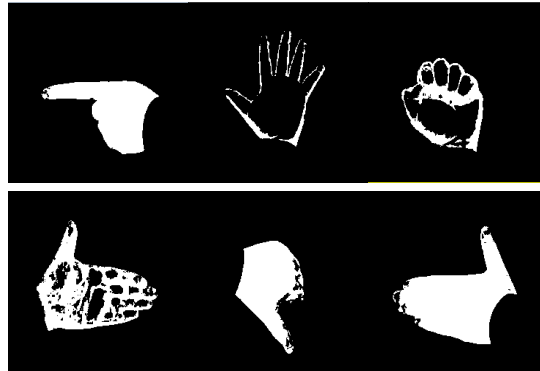
Step7: calculate S output from, if  $M = 0$  then  $S = 0$  and  $H = 180$  degrees  
if  $M \neq 0$  then  $S = (M - m) / M$

Step8: calculate H output from, if  $R = M$  then  $H = 60(b-g)$   
if  $G = M$  then  $H = 60(2+r-b)$   
if  $B = M$  then  $H = 60(4+g-r)$   
if  $H \geq 360$  then  $H = H - 360$   
if  $H < 0$  then  $H = H + 360$

Step9: output H, S, and V color space, where H in the range  $[0,360]$ , S and H in the range  $[0,100]$

We decided a range of values for each of H, S and V so it accepts the pigment of human skin, let is the input gesture colored image location  $(x, y)$ , and let  $H(x, y)$ ,  $S(x, y)$  and  $V(x, y)$  are the H, S and V bands for the HSV color space for the input gesture image at location  $(x, y)$ ; let M be the output binary image, we set  $M(x, y)$  to 1 when  $H_{min} < H(x, y) < H_{max}$  and  $S_{min} < S(x, y) < S_{max}$  and  $V_{min} < V(x, y) < V_{max}$ ; otherwise, we set  $M(x, y)$  to 0; Figure (6) is the output of applying this technique on Figure (4).

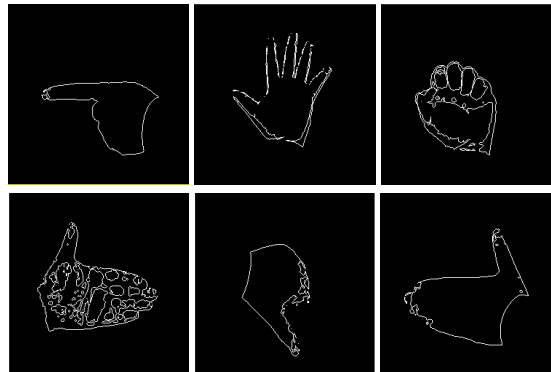
After this phase, hand boundary required to be last step in the image preprocessing, laplacian edge detector is applied as we explained before, the mask shown in Figure (7) is used and produces the output shown in Figure (8).



**FIGURE 6:** Binary Images.

0	1	0
1	- 4	1
0	1	0

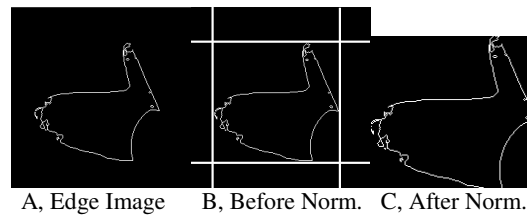
**FIGURE 7:** Laplacian Mask.



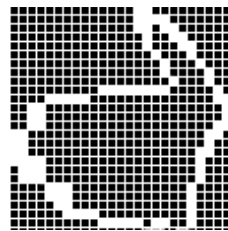
**FIGURE 8:** Edges of the Gestures.

Now, normalization operation is applied for removing the unwanted area of the gesture image, Figure (9) shows the application of normalization operation for one gesture of the samples gestures, the rest gestures are the same.

After this point, the gesture image is ready for feature extraction, as we said, the image output is 128x128 pixels, and block size is 5x5 pixels, so, each gesture image can produce 625 feature vector represents the features of this gesture, Figure (10) shown the values computed from the output gesture in Figure (9).



**FIGURE 9:** Normalization Operation.



**FIGURE 10:** Features Calculation via Dividing the Gesture.

As seen in Figure(11), out of 625 features just 98 are white and stored and the other are neglected which represents 15.68 % stored features out of 625, and 84.32 % is neglected, Equation (1) shows the mathematical implementation for the gesture division.

$$B = \sum_{k1=-2}^2 \sum_{k2=-2}^2 M(k1 + i, k2 + j) \quad (1)$$

Where i,j takes the values 2..height-3, 2.. width-3 respectively.

B: is the output brightness value.

After getting these features, and at the time of recognition during testing stage, the features of the testing gesture are calculated and compared with database features using our suggested algorithm, and the highest matching score is passed along with it is gesture and meaning, the algorithm is described below:

Consider D(i,k) is the database feature k of gesture i, and T(k) is the feature k of the testing image, M(k) is the matching status of feature k between both database gesture i and input testing gesture, for each gesture i, if D(i, k) is black area and T(k) is black area, the set M(k) as matched status, if the brightness value of D(i, k) > Threshold and T(k) > Threshold, then set M(k) as matched status, otherwise, set M(k) as non-matched status, after that the number of matched state is calculated and the recognition percentage with this database gesture is calculated via Equation (2).

$$\text{Recognition Percentage} = \frac{\text{number of matched status}}{\text{total number of blocks}} * 100 \% \quad (2)$$

And when the matching percentage exceed 85 %, the algorithm stops immediately for saving the time and matching is found, if not, the algorithm stays running until all gestures database are matched and the maximum matching percentage is passed.

### 5. TESTING AND VALIDATION

We test our system with 24 different samples, 4 gestures for each pose( total of 6 poses), and the recognition rate was 91 %, our system recognized the gestures with different transformations and achieve this percentage of recognition, the translation and scaling solved by normalization operation which produces a unified gesture image boundaries, rotation is managed here by brightness value and blocking method which allows some flexibility in gesture slope, the testing gesture is undergone in same serious of operation to get the features, after that these extracted features will be compared with the stored database features, the maximum matching score gesture is the equivalent gesture and it is meaning represents the meaning of the tested gesture, Figure (11) shows the testing gestures with their maximum matching probability.

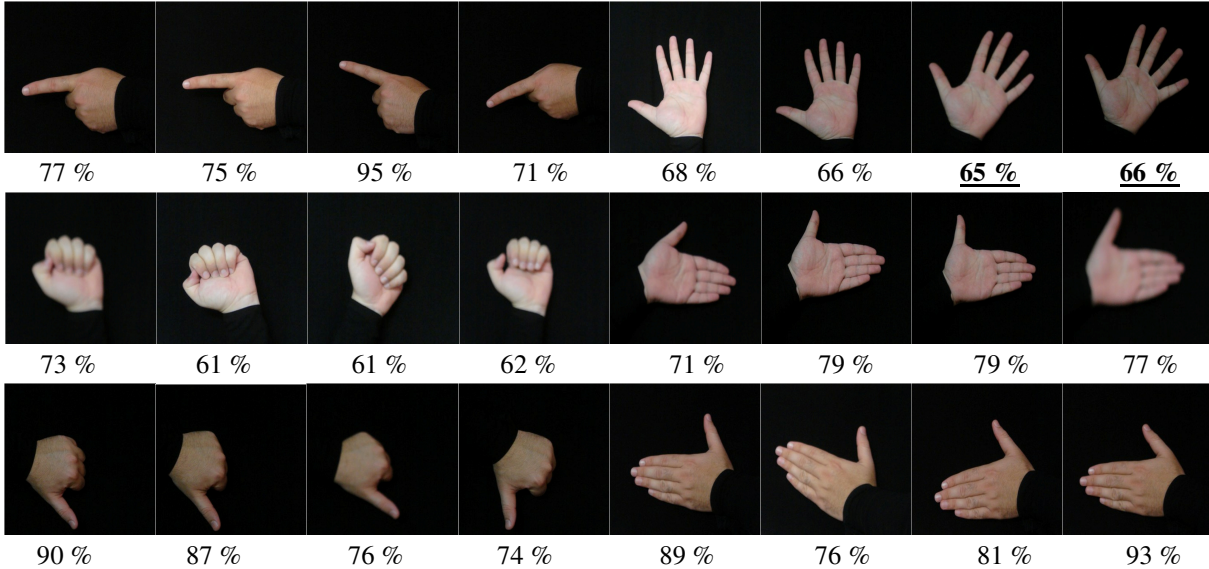


FIGURE 11: Recognition Percentage for each Tested Gesture.

In Figure (11), the underlined and bolded recognition percentage represents the non-recognized gestures by the system, they have highest probability but referring to wrong gesture. You can notice the rotation in these testing gestures and the system still recognizes these gestures. Figure (12) represents the matching chart for two selected testing gestures from Figure (11) which have matching percentage %93, %65 respectively, and first one is recognized gesture and the second one is non-recognized gesture.

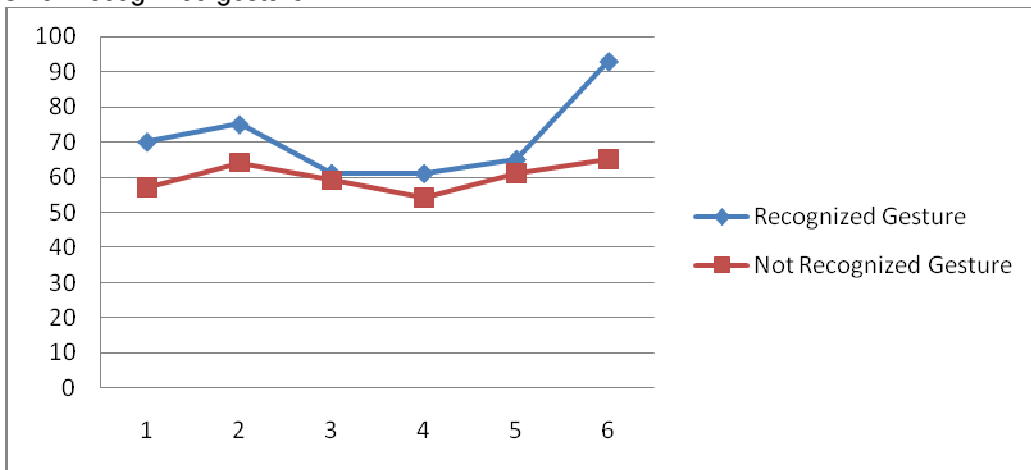
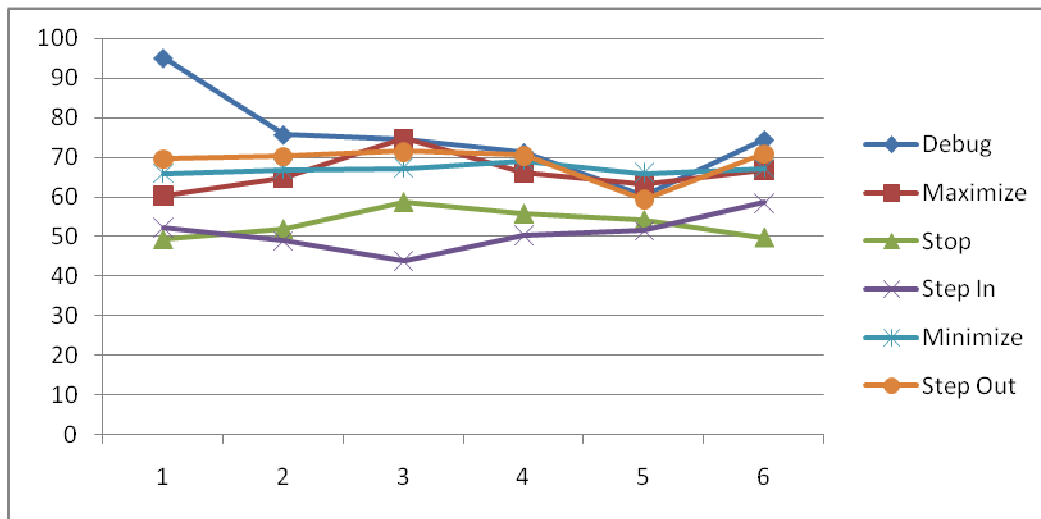


FIGURE 12: Matching Chart for Two Selected Testing Gestures.



In Figure (13), we have applied our recognition algorithm against all database feature vectors and the recognition rates are shown herein below, the aim is to reveal the prominent of the recognized gesture over all other gestures.



**FIGURE 13:** Matching Single Gesture against all Database Gestures.

## 6. CONSLUSION & FUTURE WORK

Hand gesture segmentation using HSV color model provide a promising results, this segmentation rely mainly on the pigment color of the human skin regardless the illumination which can be overcome by setting correct parameters for this model, also the seeking for skin color considered to be an efficient technique specially when the input image is cluttered with many objects, many researches commenced to enhance and develop the interaction between the human and human-made machines, template matching uses raw data and affected by rotation and transformation changes in addition to the illumination changes, in order to overcome this problem we need to increase the number of samples per gesture which indicates the increasing of the database size and in turn the processing time as well, the orientation histogram provides a promising solution for illumination changes due to the edge direction which is not affected with illumination variance, nevertheless, the hand object must dominate [6] the input gesture image and the database vocabulary must be chosen to avoid some confusion [6] in testing phase which may provide same testing result for different input gestures, by using of neural network the number of samples must be limited due to training time and the training data must be separable, in our method we have built a recognition system can handle a good degree of rotation via partitioning technique which partitions the hand gesture into blocks, and then we have applied the suggested testing algorithm.

Overall, feature selection is an important issue for gesture recognition and it is considered to be a crucial to the recognition algorithm, image preprocessing steps also important because the perfect steps can promise a unique and small feature vector which will reduce the number of samples in the database and speed up the recognition time.

In this study we have achieved 91 % recognition rate using different gestures at different rotation angles but using same conditions of illumination against uniform background, we have trained our system with 60% of gestures and tested our system with 40%, in the future work one can use the non-uniform background instead of uniform background.

## 7. REFERENCES

1. Wikipedia Internet Web Site
2. Internet Web Site. Available at:<http://commons.wikimedia.org>
3. S. Naidoo, C.W. Omlin, M. Glaser, "*Vision-Based Static Hand Gesture Recognition using Support Vector Machines*". Department of Computer Science, University of the Western Cape, South Africa, 1999
4. K. Symeonidis. "*Hand Gesture Recognition using Neural Networks*", Master Thesis, School Of Electronic And Electrical Engineering, 2000
5. R. Brunelli, T. Poggio. "*Face Recognition: Features versus Templates*". IEEE Transactions on Pattern Analysis And Machine Intelligence, 15(10):1042-1052, 1993
6. W. T. Freeman, M. Roth. "*Orientation Histograms for Hand Gesture Recognition*". Mitsubishi Electric Research Laboratories, Cambridge, Ma 02139 USA, 1994
7. X. Li. "*Gesture Recognition based on Fuzzy C-Means Clustering Algorithm*". Department Of Computer Science The University Of Tennessee Knoxville, 2003
8. X. Li. "*Vision Based Gesture Recognition System with High Accuracy*". Department of Computer Science, The University of Tennessee, Knoxville, TN 37996-3450, 2005
9. J. Wachs, U. Kartoun, H. Stern, Y. Edan. "*Real-Time Hand Gesture Telerobotic System using Fuzzy C-Means Clustering*". Department of Industrial Engineering and Management, Ben-Gurion University of the Negov, 1999
10. J. Triesch, C. Malsburg. "*Robust Classification of Hand Postures Against Complex Backgrounds*". IEEE Computer Society, In Proceedings of Second International Conference On Automatic Face And Gesture Recognition, 1996
11. T. Yang, Y. Xu. "*Hidden Markov Model for Gesture Recognition*". The Robotics Institute Carnegie Mellon University Pittsburgh, Pennsylvania 15213, 1994
12. J. J. Phu, Y. H. Tay. "*Computer Vision Based Hand Gesture Recognition using Artificial Neural Network*". Faculty of Information and Communication Technology, University Tunku Abdul Rahman (Utar), Malaysia, 2006
13. H. B. Amor, S. Ikemoto, T. Minato, H. Ishiguro. "*Learning Android Control using Growing Neural Networks*". Department Of Adaptive Machine Systems Osaka University, Osaka, Japan, 2003
14. M. Swain and D. Ballard. "*Indexing via Color Histograms*". In Proceedings of Third International Conference on Computer Vision, 390-393, 1990
15. S. Venkataraman, V. Gunaseelan. "*Hidden Markov Models in Computational Biology*". lectures in HMM
16. The AF Research Laboratory. "*Language and Cognition*". Elsevier, Neural Networks, 22: 247-257, 2009
17. H. Gunes, M. Piccardi, T. Jan. "*Face and Body Gesture Recognition for a Vision-Based Multimodal Analyzer*". Computer Vision Research Group, University of Technology, Sydney (UTS), 2007

18. Y. Lu, S. Lu, F. Fotouhi, Y. Deng, Susan J. Brown. *"A Fast Genetic K-Means Clustering Algorithm"*. Wayne State University, Kansas State University Manhattan, USA, 2000
19. B. Heisele, P. Ho, T. Poggio. *"Face Recognition with Support Vector Machines: Global versus Component-based Approach"*. Massachusetts Institute of Technology Center for Biological and Computational Learning Cambridge, 2001
20. K. Jain, R. P.W. Duin J. Mao. *"Statistical Pattern Recognition: A Review"*. IEEE Transactions on Patterns Analysis and Machine Intelligence, 22(1):4-35, 2000
21. S. S. Keerthi, O. Chapelle, D. DeCoste. *"Building Support Vector Machines with Reduced Classifier"*. Complexity, Journal of Machine Learning Research, 8:1-22, 2006
22. Abin – Roozgard. *"Convolutional Neural Networks"*. lectures in Neural Networks
23. Y. P. Lew, A. R. Ramli, S. Y. Koay, A. ali, V. Prakash. *"A Hand Segmentation Scheme using Clustering Technique in Homogeneous Background"*. Student Conference on Research and Development Proceedings, Shad Alam, Malaysia, 2002
24. C.C. Lo, S. J. Wang. *"Video Segmentation using a Histogram-Based Fuzzy C-Means Clustering Algorithm"*. Institute of Information Management, National Chiao-Tung University, Computer Standards & Interfaces, 23:429–438, 2001
25. S. Marcel, O. Bernier, J. Viallet, D. Collobert. *"Hand Gesture Recognition using Input–Output Hidden Markov Models"*. France Telecom Cnet 2 Avenue Pierre Marzin 22307 Lannion, France, 1999
26. C. Karlof, D. Wagner. *"Hidden Markov Model Cryptanalysis"*. Computer Science Division (EECS) University of California Berkeley, California 94720, 2004

## A Novel Secret Sharing Technique Using QR Code

**Jun-Chou Chuang**

*Assistant Professor*

*Department of Computer Science and communication Engineering  
Providence University 200 Chung-Chi  
Rd., Taichung, Taiwan*

lzchung@pu.edu.tw

**Yu-Chen Hu**

*Professor*

*Department of Computer Science and Information Management  
Providence University 200 Chung-Chi  
Rd., Taichung, Taiwan*

ychu@pu.edu.tw

**Hsien-Ju Ko**

*Assistant Professor*

*Department of Photonics and Communication Engineering  
Asia University 500, Liufeng Rd.,  
Wufeng, Taichung, Taiwan*

nhjko@asia.edu.tw

---

### Abstract

Any mobile device with capture function can read content from a barcode tag directly. When a barcode contains important data or privacy information, the risk of security becomes an important problem. In this paper, the QR code is employed to design the secret sharing mechanism so that the data privacy during data transmission can be enhanced. The secret data is divided into some shadows by the secret sharing mechanism and the results are embedded into barcode tags. The secret can be recovered only when the number of shadows is greater than or equal to the predefined threshold. In sum, the proposed technique improves data security for data transmission.

**Keywords:** Barcode, secret sharing, QR code.

---

### 1. INTRODUCTION

Barcode provides a convenient way [2][5][15] for people labeling a tag on a product so that people can easily and quickly identify the content of product itself. It can be classified into two types, one-dimensional (1D) barcode and two-dimensional (2D) barcode. The 1D barcodes use different width of lines and spaces to represent data, for example, code 39, code 128, EAN-13, EAN-128, ISBN, and etc. As for the 2D barcodes, they use symbol types of stacking and matrix to represent data, such as QR code [4][7][8][16][17][18], PDF417, Data Matrix, Maxi Code, and etc. Table 1 shows different types of 1D barcodes and 2D barcodes.

In generally, 1D barcodes put emphasis on “product identification” and 2D barcodes put emphasis on “product descriptions”. Because of the limitation of 1D barcode storage, only a few data like product identification is stored in 1D barcode. 2D barcodes are superior to that 1D barcode in embedding payload, error resistance, data security, and readability. In the storage size, 2D barcode can store a lot of information like product descriptions, including product

ingredient, product item, product details, web links, and etc. For error resistance, 2D barcodes can defense different levels of error occurs.

The security of 1D barcodes is lower than 2D barcodes. 1D barcodes are very easy to read by scanning the lines and the spaces. However, 2D barcodes are not easy to read a symbol pattern by human eyes. With regard to readability, 1D barcodes must scan along a single directional. If the angle of a scan line does not fit within a range, the data would not be read correctly. However, 2D barcodes get wide ranges of angles for scanning. Thus, 2D barcodes are readability.

2D Barcodes provide a unique identifier for objects and applications [1][9][10][11] [12][13][14][15] to automatic checkout system, commerce, industry, hospital, and etc. Barcodes are very convenience to automatic systems, but they have data privacy weakness. A reader device with video capture function can read the content from tags directly. When barcodes contain privacy information may result in the risk of security issue. Therefore, the confidential data is often stored in the back-end database. When a reader captures a tag, it only gets a network link from a tag and later connected to the back-end database through the Internet. A user who has access right can login database to retrieve the privacy information.

1D barcodes	<p>Code 39</p>  <p>123456</p>	<p>Code 128</p>  <p>123456</p>	<p>EAN-13</p>  <p>1 234567 890128</p>	<p>ISBN</p>  <p>9 781234 567897</p>
2D barcodes	<p>QR Code</p> 	<p>PDF417</p> 	<p>DataMatrix</p> 	<p>Maxi Code</p> 

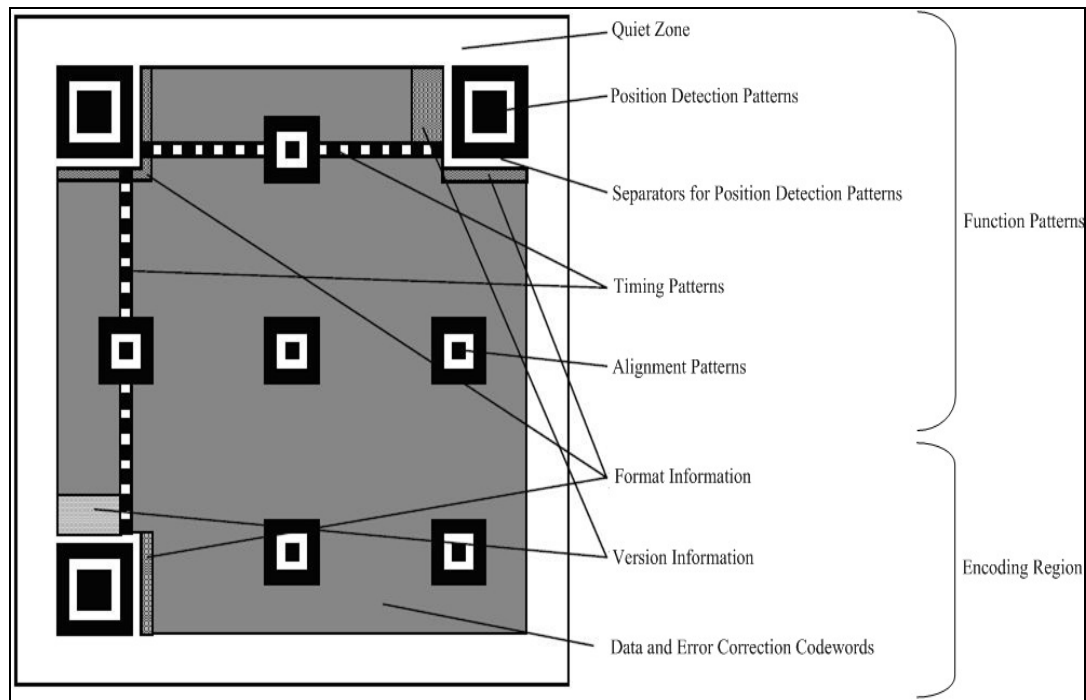
**TABLE 1:** 1D Barcodes and 2D Barcodes.

To enhance security of data privacy [3] of barcodes, we design a secret sharing technique with Quick Response code (QR code). The technique shares a confidential data into shadows and one shadow is embedded into one carrier tag. Anyone cannot recovery the original secret data from its own share. The secret can be recovered only when the number of shadows is larger than or equal to a threshold. The proposed technique does not need to connect the back-end database through Internet. Thus, the proposed technique can save much more hardware cost and can reduce the security risks transmission on the open environment.

The rest of this paper is organized as follows. In Section 2, we review the QR code. The proposed technique is described in Section 3. The security analysis and performance is listed in Section 4. Finally, the conclusions are presented in Section 5.

## 2. QR CODE

The QR code is a kind of matrix symbol, which was developed by the Japanese company Denson-Wave in 1994. Figure 1 shows the basic structure of QR code. They are quiet zone, position detection patterns, separators for position detection patterns, timing patterns, alignment patterns, format information, version information, data, and error correction codewords. They are shown in Figure 1. Some details of QR code can be refereed to [17].



**FIGURE 1:** The basic structure of QR Code.

The main features of QR code contain large capacity, small printout size, high speed scanning, advanced error correcting, and freedom direction scanning. The overall are summarized as follows.

- High data capacity: QR code can store 7,089 numeric characters and 4,296 alphanumeric characters, and 1,817 kanji characters.
- High speed scanning: A mobile phone with camera function can get the content from a barcode quickly and easily.
- Small printout size: QR Codes carry data on both horizontally and vertically, thus QR codes are better than 1D barcodes in data capacity.
- Advance error correcting: Even if 50% areas of barcode are damaged, QR codes still can be recognized correctly.
- Freedom direction scanning: The scanning direction of QR code is freedom.

### 3. DESIGN OF SECRET SHARING TECHNIQUE USING QR CODE

The proposed technique designs a secure data transmission scheme based on the secret sharing scheme with QR code. Secret sharing scheme was first proposed by Shamir in 1979 [14]. The main idea of the secret sharing scheme divides a secret into  $n$  shadows or called shares. Anyone can not decrypt the original secret from their own share. The secret can be recovered only when any of  $t$  out of  $n$  shadows ( $t <= n$ ) are hold together. The framework of the proposed scheme is listed in Figure 2.

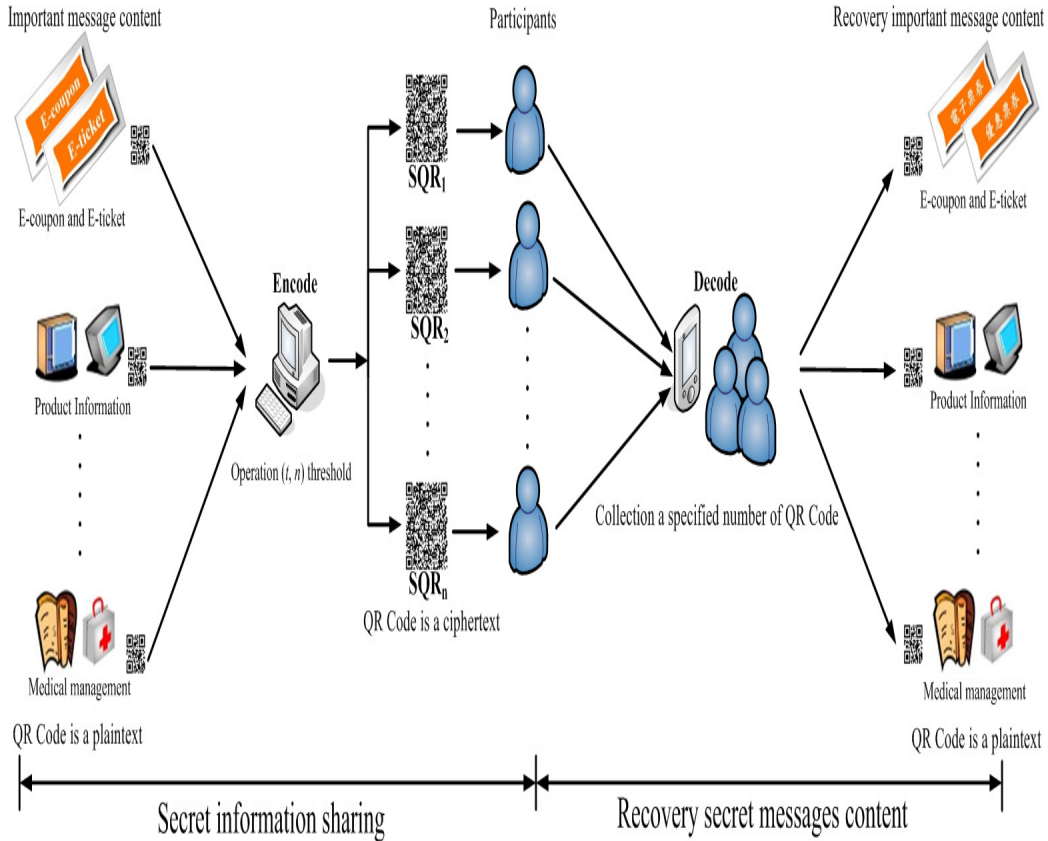


FIGURE 2: The framework of the proposed scheme.

In the proposed scheme, we first choose a value  $t$  ( $t < n$ ) and a secret key  $a_0$  and a large prime number  $p$  ( $p > a_0$ ). Next, we select  $n$  participants  $x_1, x_2, \dots, x_n$ , where  $n$  is the number of QR code tags, which used to be hidden. Next, a  $(t-1)$  degree polynomial  $f(x)$  was constructed as follows:

$$f(x) = a_0 + a_1x^1 + a_2x^2 + \dots + a_{t-1}x^{t-1} \pmod{p} \tag{1}$$

Where  $a_1, a_2, \dots, a_{t-1} \in \mathbb{Z}_p$ . So, we can generate a pair of secret share  $(x_i, f(x_i)=y_i)$  to each participant.

In the decoding procedure, anyone who gets  $t$  out of  $n$  secret shares would recovery secret data  $a_0$  by the Lagrange polynomial interpolation which was listed as belows.

$$f(x) = \sum_{a=1}^t y_{ia} \prod_{j=1, j \neq a}^t \frac{x - x_{ij}}{x_{ia} - x_{ij}} \pmod{P} \tag{2}$$

Here we illustrate an example to show how to construct secret shares. Let  $t$  value is 3,  $n$  value is 6, secret key  $a_0$  is 1234 and the prime number  $p$  is 1237. Then a  $(t-1)$  degree polynomial  $f(x)=94x^2+166x+1234$  is constructed. Assume  $x_1=1, x_2=2, x_3=3, x_4=4, x_5=5,$  and  $x_6=6$ , we can obtain six secret shares where  $f(x_1)=f(1)=1494, f(x_2)=f(2)=1942, f(x_3)=f(3)=2578, f(x_4)=f(4)=3402, f(x_5)=f(5)=4414,$  and  $f(x_6)=f(6)=5614$ . To recover the secret key  $a_0$ , we need to collect three or more secret shares. Assume we obtain three secret shares, they are  $(f(2), 1942), (f(4), 3402), (f(5), 4414)$ . Then the secret key  $a_0=1234$  can be decoded by the Lagrange polynomial interpolation as below.

$$\begin{aligned}
 f(x) &= 1942\left(\frac{x-4}{2-4} \times \frac{x-5}{2-5}\right) + 3402\left(\frac{x-2}{4-2} \times \frac{x-5}{4-5}\right) + 4414\left(\frac{x-2}{5-2} \times \frac{x-4}{5-4}\right) \\
 &= 1942 \times \left(\frac{1}{6}x^2 - \frac{3}{2}x + \frac{10}{3}\right) + 3402 \times \left(-\frac{1}{2}x^2 - \frac{7}{2}x - 5\right) + 4414 \times \left(\frac{1}{3}x^2 - 2x + \frac{8}{2}\right) \pmod{1237} \\
 &= (94x^2 + 166x + 1234) \pmod{1237}
 \end{aligned}$$

#### 4. SECURITY ANALYSIS AND PERFORMANCE

This section describes the security and the performance of the proposed scheme. The proposed scheme is based on Shamir’s secret sharing scheme. The secret data is divided into shares of shadows by secret sharing technique. The generated shadows are embedded into each QR-code tag. Anyone who wants direct to read the content from QR codes is impossible if the number of received shadows is not achieved the predefined threshold. As the result, our scheme is secure.

In Figures 3-6, the share generate and the message recovery for the (2, 3)-threshold and the (3, 3)-threshold are listed below. The plaintext is divided into three shares by secret sharing technique, and then embedding them into QR codes. The decoding procedure in the (2, 3)-threshold, the original plaintext can be recovered only if the number of the received QR codes is larger than or equal to two. As for the (3, 3)-threshold, the number of the received QR codes should be equal to three. Besides, error recovery from any of two barcode tags using (3, 3)-threshold is listed in Figure 7. Because of the number of the retrieved shares is less than the predefined threshold thus the reconstructed secret key is not correct.

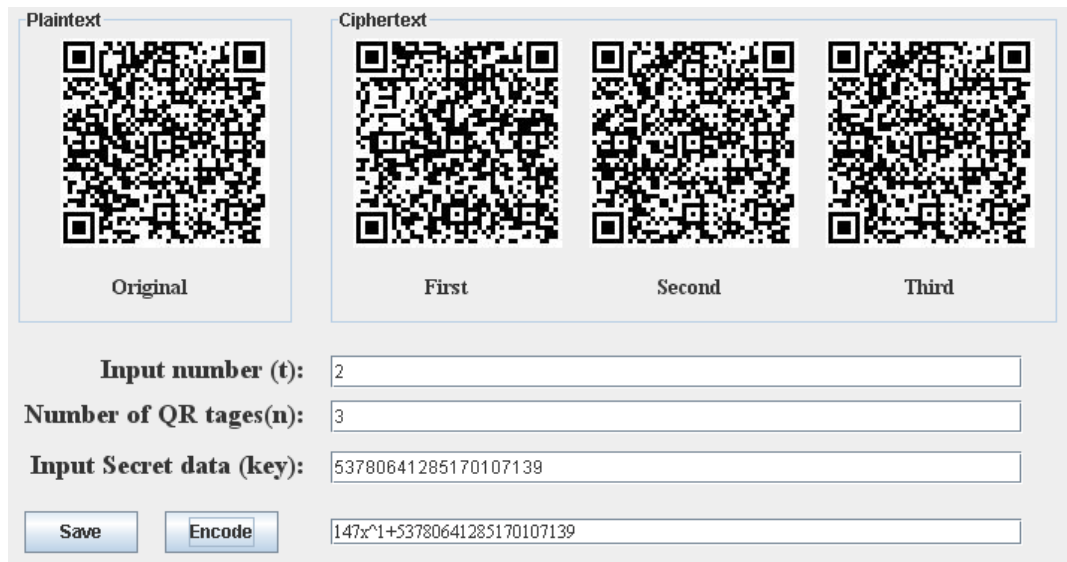


FIGURE 3: Generating three secret shares using (2, 3)-threshold.



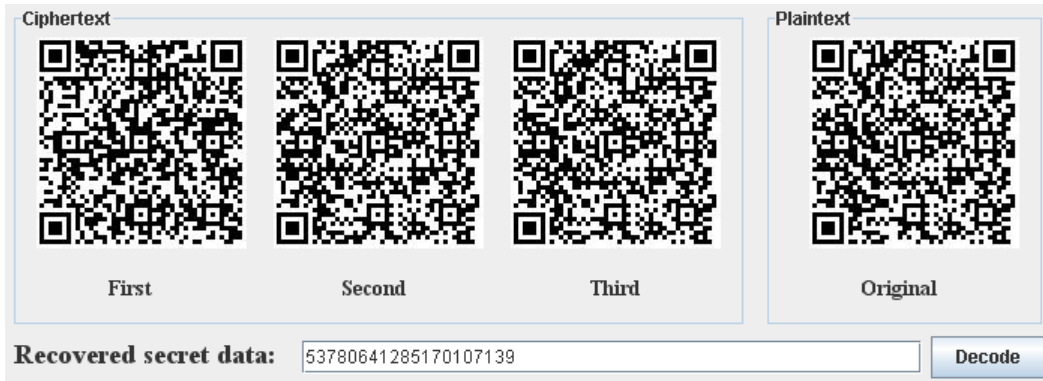


FIGURE 4: Recovery of secret message from any of two barcode tags using (2, 3)-threshold.



FIGURE 5: Generating three secret shares using (3, 3)-threshold.



FIGURE 6: Recovery of secret message from three barcode tags using (3, 3)-threshold.

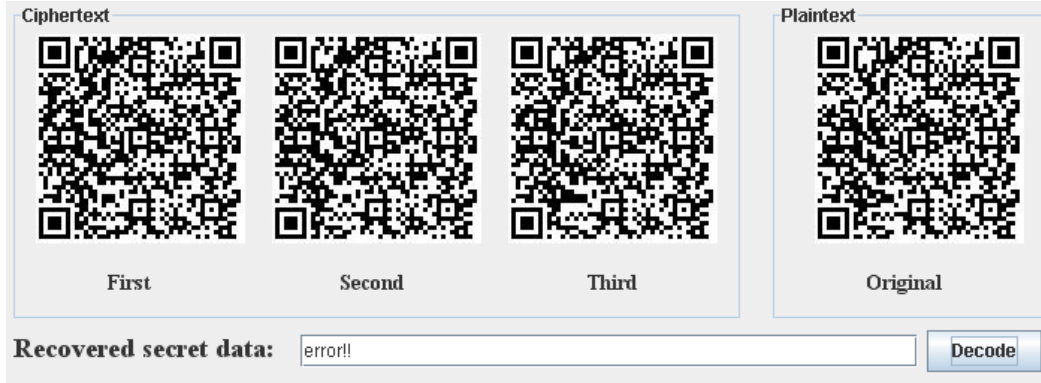


FIGURE 7: Error recovery from any of two barcode tags using (3, 3)-threshold.

As for our system performance, our system does not need to establish a remote database for QR code data searching. Meanwhile, the proposed scheme embeds the data into QR code tags directly. The cost of this paper is only tags and printing ink. Thus, the proposed scheme can save a lot of system cost.

## 5. CONCLUSIONS

In this paper, a secret sharing mechanism to enhance the security and data privacy for QR code is proposed. The proposed technique improves data security during data transmission. On the other hand, the proposed technique does not need to establish a back-end database beforehand for contents searching. It directly embeds the secret data into tags therefore the proposed technique can save a lot of hardware cost and software maintenance. The proposed technique can be applied to some applications such as electronic tickets, airline luggage inspection, medical e-health system, and others fields.

## 6. ACKNOWLEDGMENT

The authors gratefully acknowledge the support provided to this study by the National Science Council of Taiwan under grant number NSC99-2221-E-126-006.

## 7. REFERENCES

1. H. S. Al-Khalifa, "Mobile SRS: a classroom communication and assessment service". In Proceedings of the Innovations in Information Technology, United Arab Emirates, 2008.
2. T. Bouchard, M. Hemon, F. Gagnon, V. Gravel, and O. Munger, "Mobile telephones used as boarding passes: Enabling Technologies and Experimental Results". In Proceedings of the 4th Autonomic and Autonomous Systems, Gosier, Guadeloupe, 2008.
3. T. Chen, "The application of bar code forgery - proof technology in the product sales management". In Proceedings of the Intelligent Information Technology Application Workshops, Washington, DC, USA, 2008.
4. U. B. Ceipidor, C. M. Medaglia, and A. Perrone, M. D. Marsico, and G. D. Romano, "A museum mobile game for children using QR-codes". In Proceedings of the 8th International Conference on Interaction Design and Children, Italy, 2009.
5. Y. J. Chang, S. K. Tsai, and T. Y. Wang, "A context aware handheld wayfinding system for individuals with cognitive impairments". In Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility, Halifax, Nova Scotia, Canada, 2008.

6. N. Fujimura and M. Doi, "*Collecting students' degree of comprehension with mobile phones*". In Proceedings of the 34th Annual ACM SIGUCCS Conference on User Services, Canada, 2006.
7. T. Falas and H. Kashani, "*Two-dimensional barcode decoding with camera-equipped mobile phones*". In Proceedings of the Pervasive Computing and Communications Workshops, White Plains, NY, USA, 2007.
8. J. Z. Gao, L. Prakash, and R. Jagatesan, "*Understanding 2D-barcode technology and applications in m-commerce – design and implementation of a 2D barcode processing solution*". In Proceedings of the Computer Software and Applications Conference, Beijing, China, 2007.
9. T. Kamina, T. Aoki, Y. Eto, N. Koshizuka, J. Yamada, and K. Sakamura, "*Verifying identifier-authenticity in ubiquitous computing environment*". In Proceedings of the Advanced Information Networking and Applications Workshops, Ontario, Canada, 2007.
10. B. Lingyan, F. Zewei, L. Min, and W. Weining, "*Design and implementation of the airline luggage inspection system base on link structure of QR code*". In Proceedings of the Electronic Commerce and Security, Guangzhou, 2008.
11. T. Y. Liu and Y. L. Chu "*Handheld augmented reality supported immersive ubiquitous learning system*". In Proceedings of the Systems, Man and Cybernetics, Singapore, 2008.
12. J. Rouillard, "*Contextual QR codes*". In Proceedings of the 3rd International Multi-Conference on Computing in the Global Information Technology, Athens, Greece, 2008.
13. S. Reiser and R. Bruce, "*Service learning meets mobile computing*". In Proceedings of the Annual Southeast Regional Conference, Auburn, Alabama, 2008.
14. A. Shamir, "*How to Share a Secret*". Communication of the ACM, 22(11): 612-613, 1979.
15. G. Starnberger, L. Frohofer, and K. M. Goeschka, "*QR-TAN: Secure mobile transaction authentication*". In Proceedings of the Availability, Reliability and Security, Fukuoka, Japan, 2009.
16. Y. L. Yeh, J. C. You, and G. J. Jong, "*The 2D bar-code technology applications in medical information management*". In Proceedings of the Intelligent Systems Design and Applications, Kaohsiung, Taiwan, 2008.
17. ISO/IEC18004, "*Information technology-automatic identification and data capture techniques*". Bar Code Symbology - QR Code.
18. Denso-wave, <http://www.denso-wave.com/qrcode/index-e.html>.

## Automated Monitoring System for Fall Detection in the Elderly

**Shadi Khawandi**  
*University of Angers*  
*Angers, 49000, France*

skhawandi@hotmail.com

**Bassam Daya**  
*Lebanese University*  
*Saida, 813, Lebanon*

**Pierre Chauvet**  
*University of Angers*  
*Angers, 49000, France*

---

### Abstract

Falls are a major problem for the elderly people living independently. According to the World Health Organization, falls and sustained injuries are the third cause of chronic disability. In the last years there have been many commercial solutions aimed at automatic and non automatic detection of falls like the social alarm (wrist watch with a button that is activated by the subject in case of a fall event), and the wearable fall detectors that are based on combinations of accelerometers and tilt sensors. Critical problems are associated with those solutions like button is often unreachable after the fall, wearable devices produce many false alarms and old people tend to forget wearing them frequently. To solve these problems, we propose an automated monitoring that will detects the face of the person, extract features such as speed and determines if a human fall has occurred. An alarm is triggered immediately upon detection of a fall.

**Keywords:** Fall Detection; Monitoring System; Face Detection; Elderly; Webcam.

---

### 1. INTRODUCTION

Falls and sustained injuries among the elderly are a major problem worldwide. The proportion of people sustaining at least one fall during one year varies from 28-35% for the age of 65 and over. The risk of falling increases with age, and in 2 cases out of 3 it happens at home. People that experience a fall event at home, and remain on the ground for an hour or more, usually die within 6 months. In the last years there have been many commercial solutions aimed at automatic and non automatic detection of falls [1][2][3][4] like the social alarm, which is a wrist watch with a button that is activated by the subject in case of a fall event, and the wearable fall detectors that are based on combinations of accelerometers and tilt sensors. Critical problems are associated with those solutions; the main problem in the first solution is that the button is often unreachable after the fall especially when the person is panicked, confused, or unconscious. In the second solution, the problem is that these devices produce many false alarms, and old people tend to forget wearing them frequently [5][6][7]. To solve these problems, we propose a framework for the monitoring of elderly people, including fall detection capabilities, by using a webcam system without any user intervention. Cameras in in-home assistive system present several advantages over different sensors: they are less intrusive because installed on building (not worn by users), they are able to detect multiple events simultaneously and the recorded video can be used for post verification and analysis.

This paper presents an automated monitoring system based on image processing of the person's movement in real-time; this system will detect the face of the person, extract features such as speed and determine if a human fall has occurred. An alarm is triggered immediately upon detection of a fall.

In this paper, we present an automated monitoring system based on image processing in real time; this system detects the face of a person in a given area, collects data such as the speed of movement of the person, and determines whether the person has suffered a fall; an alarm is triggered immediately upon the detection of a fall.

## 2. WEBCAM MONITORING SYSTEM

This section describes the monitoring system, as well as its main functions. The system detects the face, extracts features (speeds, distance face-webcam, face presences), determines if a fall occurred and triggers an alarm upon fall detection.

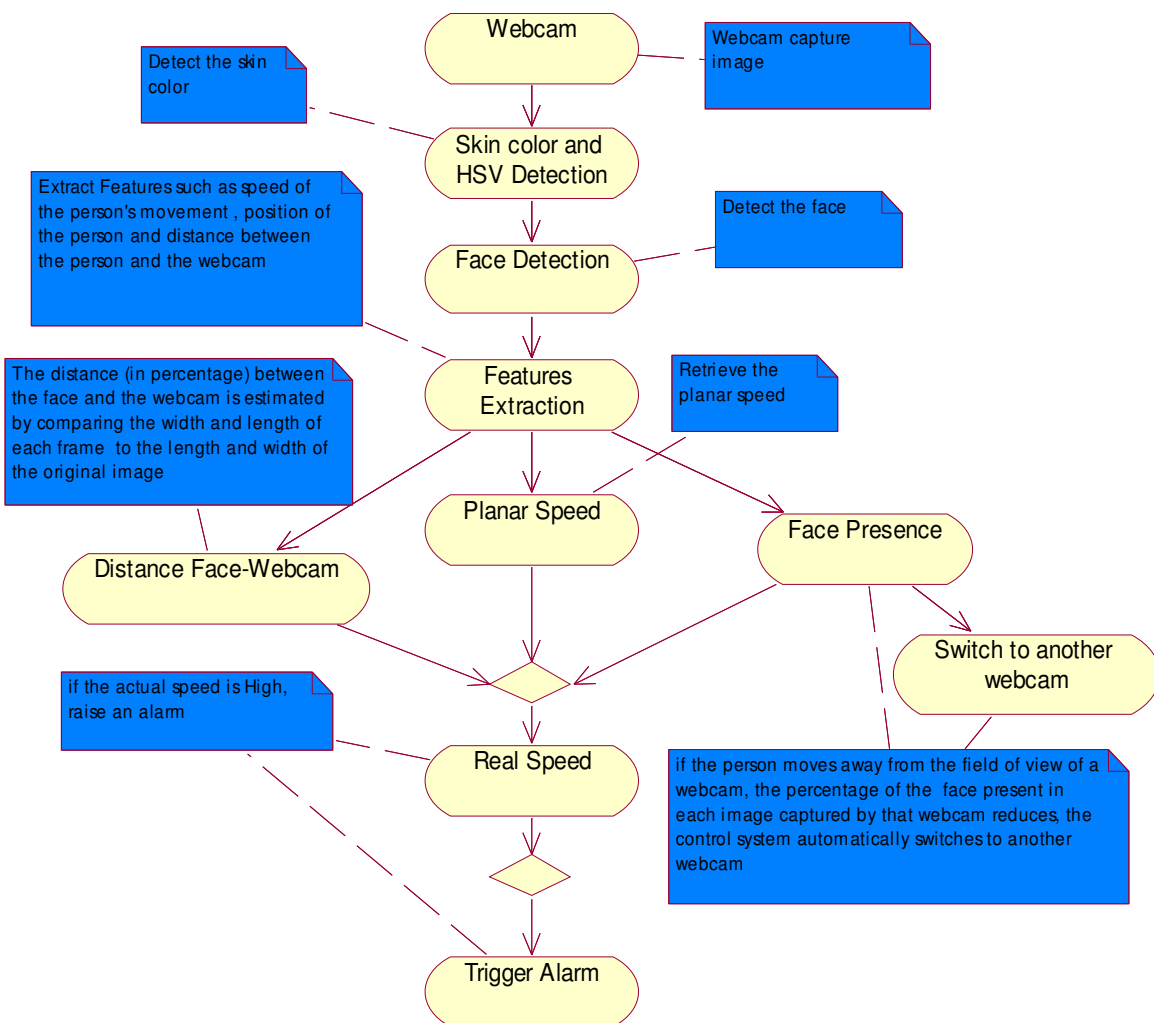


FIGURE 1: Monitoring System Architecture

## 2.1 Webcam

The Webcam is the entry point of the monitoring system. It detects the person's movement. Notes that we will have more than one webcam in order to cover all the area. Based on the presence of the face according to the webcam, a switch will be done to the other webcam. So at this phase, we have images of the person's movement that will be processed later.

## 2.2 Skin color and HSV detection

After having images of the person caught by the webcam, the system will proceed to the detection of skin color, which is an effective way often used to define a set of areas that could contain a face or hands. In addition, it can be accomplished easily based on thresholds in an appropriate color space. Different color spaces are used as skin color pixels: RGB, RGB standard, HSV (or HSI), YCrCb ... HSV is used which provides very good results. After detection of skin color, image filtering will be applied (erosion and dilation).

## 2.3 Face detection

After identifying skin areas, it is also necessary to distinguish the face, for this, the correlation is done between the detected object and the ellipse. The correlation technique is very effective and efficient. This technique is based on an approach using the comparison with an ellipse.

## 2.4 Features Extraction

After detecting the face of the person, features properties are needed to be extracted (speed, position, and distance from webcam) in order to determine if a human fall has occurred.

### 2.4.1 Distance Face-Webcam

After extracting the properties of the detected face, we can determine the distance face-webcam. By comparing the width and length of each frame to the length and width of the original image we obtain the percentage of this distance.

First, knowing the width and length of the first frame, we calculate the area (Area = length \* width) in order to know the distance (%) in the initial image (we assume 50,000 pixel as maximum area). Secondly, we determine the following:

- $X = \text{width} / \text{width}_0$   
where width is the width of the detected face for each frame  
Width<sub>0</sub> is the width of the detected face for the initial frame
- $Y = \text{length} / \text{length}_0$   
where length is the length of the detected face for each frame  
Length<sub>0</sub> is the length of the detected face for the initial frame

Third, according to the value of X or Y, we obtain the percentage of the distance for the current frame.

The purpose is to take the nearest value of 1; if the person detected by the camera turned his face to the right / left or up / down, only one parameter will change: if the move is left / right, the width varies from initial width (width<sub>0</sub>) while the length is very close to the initial length (length<sub>0</sub>). If the face moves up or down, in this case the length varies and the width is close to initial width.

### 2.4.2 Face Presence

For the fall detection and surveillance of an elderly person, we should use more than one webcam as this person may change his position to become in an area not covered by the webcam. To handle this problem, the percentage of the presence of the face for a webcam should be calculated. So if the person moves, the control automatically turns the webcams depending on the percentage of the presence of the face calculated for each webcam. This percentage is obtained by using the following expression:

$$Z = X/Y = (\text{width}/\text{width0}) * (\text{length0} / \text{length})$$

### 2.4.3 Planar Speed and Real Speed

The planar speed is calculated based on the following formula:

$$\text{Speed} = \text{distance}/\text{time} \quad (\text{pixel}/\text{sec})$$

Distance: is distance between two faces in consecutive frames (pixel)

Time: processing time between two consecutive images

The fall detection depends on the speed and the distance face-webcam at the same time. We therefore define the real speed as:

$$\text{Real speed} = \text{planar speed} * (1/Z) \quad (Z \text{ is defined in section 4.2})$$

Having the real speed, we can adopt the classification below:

- Low speed : 0-200
- Normal speed : 200-400
- Medium speed: 400-600
- High speed: 600 and up

## 3. RESULTS

In this section, we present the results obtained using the proposed monitoring system.

### 3.1 Distance between face and webcam

Frame	Width	Length	X = width/width0	Y = length/length0	Zoom (zoom %)	Planar Speed (pixel/s)	Actual speed
1	113	159	1	1	35.93 %	-	
5	104	148	0.920	0.930	33.44 %	6.3434	18.97
7	94	133	0.831	0.836	30.05 %	21.211	70.59
9	81	114	0.716	0.716	25.76 %	0	0
15	56	81	0.495	0.509	18.30 %	6.2192	33.98
16	55	78	0.486	0.490	17.62 %	8.6235	48.94

**TABLE 1:** Planar and actual speeds calculated on basis of distance between face and webcam

In the above table, frames 5 and 15 have almost the same planar speed but different zoom (z) values. Therefore, the actual speeds are different.

### 3.2 Face presence

By considering 6 frames detected by the webcam, we obtained the following results:

Frame	Width	Length	X = width/width0	Y = length/length0	Zoom (zoom %)	Face presence (%)	Remarks
1	157	203	1	1	63.74 %	100 %	
4	131	193	0.834	0.950	60.59 %	87.76 %	
7	73	172	0.421	0.847	54.01 %	54.88 %	Switch to another webcam
10	101	189	0.64	0.931	69.34 %	69.09 %	
14	132	192	0.840	0.945	60.29 %	88.89 %	
16	140	198	0.891	0.975	62.17 %	91.42 %	

**TABLE 2:** Face presence



### 3.3 Fall detection

#### 3.3.1 Results With 8 frames

By considering 8 frames detected by the webcam, we obtained the following results:

Frame	Width	Length	X = width/ width0	Y = length /lengt h0	Zoom (zoom %)	Planar Speed (pixel/s)	Actual speed
1	84	85	1	1	14.28 %	-	
2	102	91	1.214	1.070	15.28 %	80.67	527.97 Medium
3	83	90	0.988	1.058	15.12 %	63.76	421.72 Medium
4	82	91	0.976	1.070	15.29 %	40.32	263.72 Normal
5	75	86	0.982	1.011	14.45 %	85.85	594.15 Medium
6	80	76	0.952	0.894	13.60 %	29.57	217.43 Normal
7	81	89	0.964	1.047	14.95 %	57.80	386.64 Medium
8	80	85	0.952	1	14.28 %	87.83	615.10 High ALARM

**TABLE 3:** Fall detection with 8 frames

### 3.3.2 Results With 5 frames

By considering 5 frames detected by the webcam, we obtained the following results:

Frame	Width	Length	Surface	X = width/width0	Y = length/length0	Zoom (zoom %)	Planar Speed (pixel/s)	% of presence	Actual speed
1	82	90	7380	1	1	14.7 %	-	100 %	
2	83	85	7055	1.01	0.94	14.9 %	13.3	100 %	89.24 Low
3	94	99	9306	1.14	1.10	16.2 %	58.2	100 %	358.5 Normal
4	83	98	8134	1.01	1.08	14.9 %	32.0	92.96 %	214.8 Normal
5	88	85	7480	1.07	0.94	15.8 %	64.0	100 %	404.3 Medium

**TABLE 4:** Fall detection with 5 frames

From the table above, for frames 2 and 4, we have two very different areas with same zoom (effect of turning left or right, in this case the percentage presence will be very different).

### 3.4 About the results

As mentioned in the tables above, having the distance between the face and the webcam, with the face presence and the planar speed, we were able to retrieve the actual speed and detect a fall. All the calculations are based on the formulas/expressions mentioned in the previous sections.

## 4. DISCUSSION, CONCLUSION & FUTURE WORK

Elderly people represent the fastest growing segment of the population around the world. The public health services institutions will have budget problems with more pressure to limit costs. In addition, Elderly people are not used to new technologies and electronic services, and their acquisition process is always more difficult than in young people. This will justify the demand on services oriented to elderly people to help them live longer in their home increasing their quality of life. The first step in the design of this system has been to clearly identify the elderly people's requirements and the lack in the existing products to meet these requirements. The most important requirements implemented by the proposed monitoring system are:

### 4.1 Reliability of the system

One of the main aims of the monitoring system is to give more confidence to elderly people in their daily life, and such objective will be achieved with such highly reliable system. Elderly people do not want false detections neither they want falls not detected.

#### 4.2 Functionalities and services to be implemented

A wide offer of services can be implemented in such a platform. However, according to elderly people, the most important service for them to be implemented is the automatic fall detection without the need to push a button to raise the alarm. The automatic detection is continuously real time carried out while the elderly person performs any activity in its daily life. The implemented algorithm detects when a fall occurs based on fall patterns. The monitoring system is able to differentiate daily activity from falls, and so to avoid false alarms. The activity monitoring classifies the activity according to 4 levels: low, normal, medium and high level activity. An alarm will be raised when the system detects high level activities.

The proposed monitoring system offers complete activity monitoring and automatic fall detection and allows detecting distress situation and triggering an alarm upon fall detection. Tests and results show that the system is reliable and meets well the requirements. The proposed system in the paper can be treated as a strategy to be followed for any other future work in the same domain. Such system has a lot of applications; it reduces the hospitalization costs and improves life conditions for elderly peoples.

#### 5. REFERENCES

1. N. Noury, T. Hervé, V. Rialle, G. Virone, E. Mercier. "Monitoring behavior in home using a smart fall sensor and position sensors". In IEEE-EMBS. Microtechnologies in Medicine & Biology", Lyon-France, Oct 2000; 607-610.
2. N. Noury, A. Fleury, P. Rumeau, A.K. Bourke, G. O. Laighin, V. Rialle, J.E. Lundy, "Fall Detection - Principles and Methods," 29th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society, Lion (France), pp. 1663-1666, August 2007.
3. N. Noury et al. "A telematic system tool for home health care ". Int. Conf. IEEE-EMBS, Paris, 1992; part 3/7, pp. 1175-1177.
4. B.G. Celler et al. " An instrumentation system for the remote monitoring of changes in functional health status of the elderly ". In, 2, N. F. Sheppard, et al., eds., Int. Conf. IEEE-EMBS, New York, 1994, pp. 908-909.
5. G. Coyle et al." Home telecare for the elderly "Journ. Of telemedicine and telecare 1995, 1, pp. 1183-1184.
6. G. Williams et al. "A smart fall and activity monitor for telecare application ". Int. Conf. IEEE-EMBS, Hong-Kong, 1998, pp.1151-1154 Conference on, 13(16):493 – 498, 2008.
7. Yamaguchi. "Monitoring behavior in home using positioning sensors" Int. Conf. IEEE-EMBS, Hong-Kong, 1998; 1977-79

## Image Fusion and Image Quality Assessment of Fused Images

**Manjusha Deshmukh**

*Saraswati college of Engg.,  
Navi Mumbai,  
410210, India,*

manju0810@yahoo.com

**Udhav Bhosale**

*Rajiv Gandhi Institute of Technology,  
Mumbai,  
India*

udhav.bhosle@gmail.com

---

### Abstract

Accurate diagnosis of tumor extent is important in radiotherapy. This paper presents the use of image fusion of PET and MRI image. Multi-sensor image fusion is the process of combining information from two or more images into a single image. The resulting image contains more information as compared to individual images. PET delivers high-resolution molecular imaging with a resolution down to 2.5 mm full width at half maximum (FWHM), which allows us to observe the brain's molecular changes using the specific reporter genes and probes. On the other hand, the 7.0 T-MRI, with sub-millimeter resolution images of the cortical areas down to 250  $\mu\text{m}$ , allows us to visualize the fine details of the brainstem areas as well as the many cortical and sub-cortical areas. The PET-MRI fusion imaging system provides complete information on neurological diseases as well as cognitive neurosciences. The paper presents PCA based image fusion and also focuses on image fusion algorithm based on wavelet transform to improve resolution of the images in which two images to be fused are firstly decomposed into sub-images with different frequency and then the information fusion is performed and finally these sub-images are reconstructed into result image with plentiful information. . We also propose image fusion in Radon space. This paper presents assessment of image fusion by measuring the quantity of enhanced information in fused images. We use entropy, mean, standard deviation and Fusion Mutual Information, cross correlation, Mutual Information Root Mean Square Error, Universal Image Quality Index and Relative shift in mean to compare fused image quality. Comparative evaluation of fused images is a critical step to evaluate the relative performance of different image fusion algorithms. In this paper, we also propose image quality metric based on the human vision system (HVS).

**Keywords:** Hotelling Transform, Wavelet Transform, Radon Transform, Image Registration, Image Fusion.

---

### 1. INTRODUCTION

Image fusion is useful technique for merging similar sensor and multi-sensor images to enhance the information. Modern imaging technologies visualize different aspects of disease in a non-invasive way. Considerable progress has been made in the fusion of images from different

imaging modalities using software approaches. One goal of fusion software is to align anatomical and functional images and allow improved spatial localization of abnormalities. Image fusion takes place at three different levels i.e. pixel, feature and decision. Image fusion methods can be broadly classified into two that is special domain fusion and transform domain fusion. Averaging, Brovery method, Principal Component Analysis (PCA), based methods are special domain methods. But special domain methods produce special distortion in the fused image .This problem can be solved by transform domain approach. The multi-resolution analysis has become a very useful tool for analyzing images. The discrete wavelet transform has become a very useful tool for fusion. The images used in image fusion should already be registered. Mis-registration is a major source of error in image fusion. Pixel level fusion technique is used to increase the special resolution of the multi-spectral image. Application of image fusion include improving geometric correction, enhancing certain features not visible in either of the single data alone, change detection using temporal data sets and enhancing provide a complete information for diagnosis.

Image fusion needs image registration. Choice of method of image registration depends on application. Goal of image registration is to find a transformation that aligns one image to another. In image registration, one dataset is regarded as the reference data and other as sensed data. Sensed data is matched relative to the reference data. A large number of automatic image registration methods have been proposed and surveys can be found in [1], [2], [ 3]. Image registration at a very basic level can be studied from [4].

P.A. Vanden Elsen et al. proposed that a single composite image from different modality images of the same subject and provide a complete information for diagnosis.[11]. H. Li, B.S. Manjunath and S. K. Mitra adopted wavelet transform for multisensor image fusion.[12]. Researchers also proposed that wavelet based fusion method retains and inherits the main merits of tower shaped transformation [13, 14]. David A.Y. proposed method for image merging by means of discrete two dimensional wavelet transform [15]. Mallat and Zhong proposed that if the wavelet coefficients undergo a modification like coefficient merging, quantization etc. then the inverse transform preserves this modification because the transform is non redundant [16]. Sveinsson et al. proposed cluster based feature extraction and data fusion in the wavelet domain [17]. Gorzelli explained possibilities and limitations to use wavelets in image fusion [18]. Lau Wai Leung et al compared image fusion techniques using entropy and image noise index (INI) [19]. Chavez et al. proposed three different methods to merge multiresolution and multispectral data [20]. 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 [5]. 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.[9]. Wen Cao and Bicheng proposed PCAT (Principal Component Analysis Transform) and WPT (Wavelet Packet Transform) for remotely sensed image fusion [6]. Jiangsheng You, Weiguo Lu, Jian Li et. al. presented use of Radon transform for image matching [19]. Lau Wai Leung , Bruce King and Vijay Vohora compared image fusion techniques using entropy and INI [20]. Rockinger, O., proposed new merger based on shift invariant discrete wavelet transform(SIDWT) theory using maximum value selection rule of approximation coefficients for landslide characteristic enhancement [21]. Ramac, L. C., Uner, M. K., Varshney, P. K., presented Morphological filters and wavelet based image fusion. [22]. Nunez, J., proposed Multiresolution based image fusion with additive wavelet decomposition [23]. Alexander Toet, van Ruyven, J.J. & Valeton, J.M. introduced a hierarchical image merging scheme based on a multiresolution contrast decomposition i.e. the ratio of a low-pass pyramid[24]. Vivek Maik , Jeongho Shin and Joonki Paik Presented a pattern selective fusion method which provides a mechanism for combining multiple monochromatic images through identifying salient features in source images and combining those features into a single fused image[25]. Wen Doua, Yunhao Chen presented the relationships between image fusion methods aiming to reveal the nature of various methods[26]. Andrew P. Bradley proposed number of approaches to reducing, or removing, the problem of shift variance in the discrete wavelet transform (DWT) . They proposed over complete DWT (OCDWT) [27]. Milad Ghantous, Soumik Ghosh and Magdy Bayoumi

presented a hybrid image fusion scheme that combines features of pixel and region based fusion, to be integrated in a surveillance system [28]. Toet J. Ruvan and J. Valeton proposed image merging by contrast pyramid [29]. Alexander Toet presents a scheme to enhance image contrast by nonlinear multiplication of successive layers of the ROLP image decomposition [30]. P. Burt, E. Adelson proposed a Laplacian pyramid based approach for image fusion [31]. Xydeas, C., and Petrovic, V. assess pixel-level image fusion algorithms in their paper. [32].

This paper is organized as follows literature survey has been presented in section I, section II deals with fusion algorithms. Fused image evaluation methods have been presented in section III. Result is presented in section IV, We conclude in section V and bibliography has been presented in section VI.

## 2. FUSION ALGORITHMS

The details of PCA, Wavelet algorithm, Radon algorithm and their use in image fusion as well as simple average fusion algorithm are explained in this section. Experiments are carried out on following sample image.

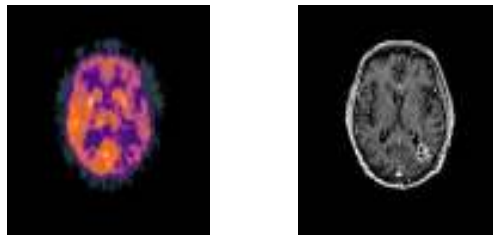


FIGURE 1: Sample images , PET (left ) and MRI (right).

### 2.1 Principal Component Analysis

Principal component analysis is one of the most frequently used dimension reduction method. Principal component analysis also called as Hotelling Transform [15]. 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, for eg. , 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 follows

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{---} \quad \langle 1 \rangle$$

Where,  $x$  is column vector. This one vector represents one common pixel in all two images. If images are of size  $P \times Q$ , there will be total of  $N=P*Q$  such  $n$ - dimensional vectors. We can assume these vectors as random quantities, Mean vector of population is

$$m_x = E[x] \quad \text{---} \quad \langle 2 \rangle$$

The covariance matrix of vector population is

$$C_x = E\{(x - m_x)(x - m_x)^T\} \quad \text{---} \quad \langle 3 \rangle$$

Element  $C_{ij}$  of  $C_x$  is the variance of  $x_i$ , the  $i^{\text{th}}$  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

$$m_x = \left(\frac{1}{N}\right) \sum_{k=1}^N x_k \text{ --- (4)}$$

Thus covariance matrix can be estimated as

$$C_x = \left(\frac{1}{N}\right) \sum_{k=1}^N x_k x_k^T - m_x m_x^T \text{ --- (5)}$$

Since  $C_x$  is real and symmetric, it is possible to find a set of N ortho-normal eigenvectors. Let  $e_i$  and  $\lambda_i$  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 correspond its smallest eigenvalue. If we use A as transformation matrix to map the x's into y. Then y is given by

$$Y = A(x - m_x) \text{ --- (6)}$$

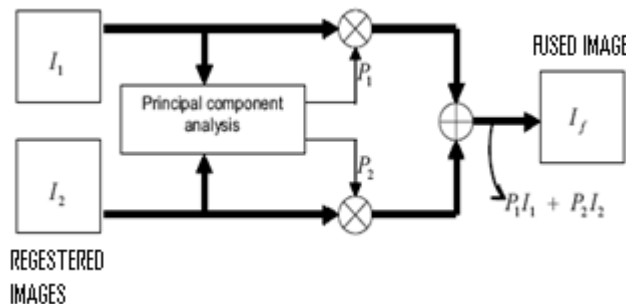
Above expression of 'y' is called Hotelling Transform or Principal Component Transform . Alternatively, the Hotelling Transform can be viewed as the discrete version of the Karbunen-Loeve transform (KLT)[4].

**A. PCA Algorithm**

The most straightforward way to build a fused image of several input images is performing the fusion as a weighted superposition of all input images. The optimal weighting coefficients, with respect to information content and redundancy removal, can be determined by a principal component analysis (PCA) of all input intensities. By performing a PCA of the covariance matrix of input intensities, the weightings for each input image are obtained from the eigenvector corresponding to the largest eigenvalue.

Arrange source images in two-column vector.

- Organize the data into column vector. Let S is the resulting column vector of dimension 2 X n.
- Compute empirical mean along each column. The empirical mean vector Me has a dimension 1 X 2.
- Subtract Me from each column of S. The resulting matrix X has dimension 2 X n.
- Find covariance matrix C of matrix X. Mean expectation will be equal to covariance of X.
- Compute eigenvectors and eigenvalue and sort them by decreasing eigenvalue.
- Consider first column of eigenvector which correspond to larger eigenvalue to compute normalized component  $P_1$  and  $P_2$ .





**FIGURE 2:** Image fusion scheme employing PCA and fused image

**Image Fusion based on Wavelet Decomposition**

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

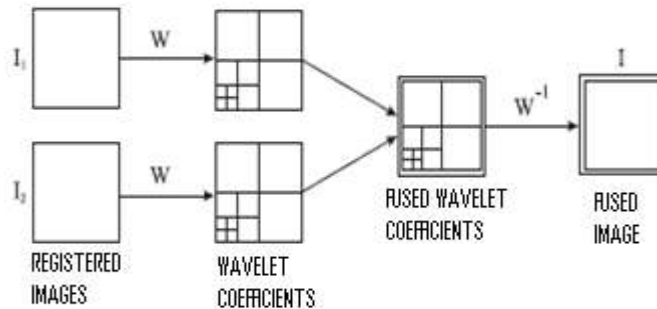
$$f(t) = \sum_{j,k} a_{j,k} \psi_{jk}(t) \dots \dots \dots \langle 7 \rangle$$

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 obtaine by applying DWT across rows and columns of an image.The two dimensional DWT of image  $f(x,y)$ is

$$f(x,y) = \sum_{j,k} C_{J_0}(k,l) \phi_{j,k,l}(x,y) + \sum_{S=H,V,D} \sum_{J=J_0}^{\infty} \sum_{k,l} D_j^S[k,l] \psi_{j,k,l}^S(x,y) \dots \dots \dots \langle 8 \rangle$$

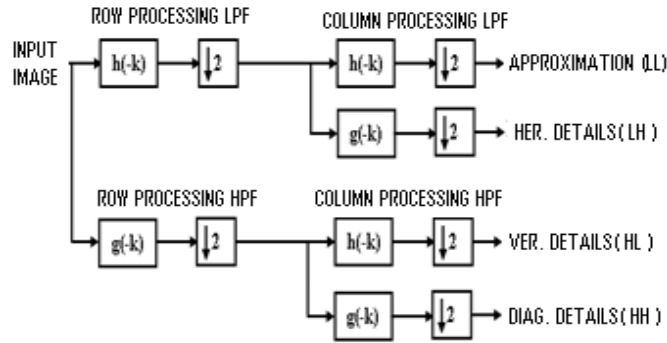
Where  $C_{J_0}$  is approximation coefficient,  $\phi_{j,k,l}(x,y)$  is scaling function,  $D_j^S$  is set of detail coefficients and  $\psi_{j,k,l}$  is set of wavelet function

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 3 and dyadic representation of the DWT is shown in figure 4 . 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.



**FIGURE3:** Wavelet multi-dimensional fusion



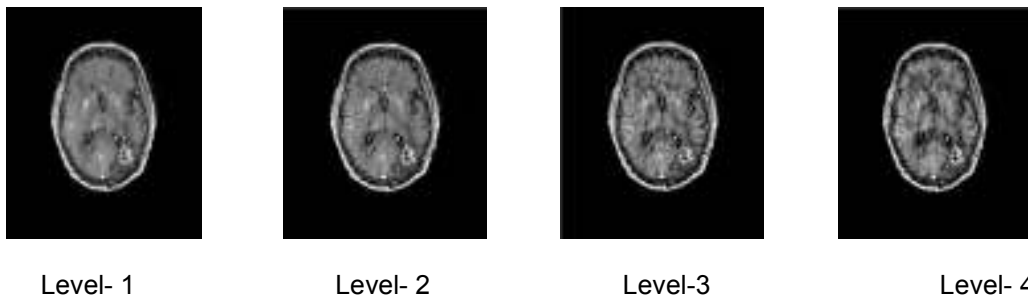


**FIGURE 4:** Two-dimensional orthogonal wavelet decomposition

**A. Wavelet based algorithm**

- 1) Apply wavelet transformation separately to each source image to establish various images of wavelet tower shaped transformation.
- 2) Fuse images at each transformation level.
- 3) Apply inverse Wavelet transform on fused wavelet pyramid.

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 Mallet's theories. The most important step for fusion is the formation of fusion pyramid. It is difficult to decide a uniform standard for fusion principle. We applied the method to PET-MRI head image. We used mutual information based method for registering source images. In the process of fusion, we fused images at four different levels. In the next section, we make a quantitative evaluation of fusion at different levels.



**FIGURE 5:** Wavelet based fused images at different levels

**Simple Average based Image Fusion**

This is a very basic technique of image fusion. Image fusion could be achieved by simple averaging corresponding pixels in each input image as follows

$$I_f(x, y) = \frac{I_1(x, y) + I_2(x, y)}{2} \dots \dots \dots (9)$$



FIGURE 6: Image fused by Averaging

### Image fusion based on Radon Transform

Radon transform is used to find linear features. Now a days, Houghs transform, Trace transform and Radon transform received much attention of researchers. Houghs Transform is a derivative of Radon transform and Radon transform is a special case of trace transform [24]. These three transforms are able to transform two dimensional images with lines into a domain of line parameters where each line in the image give a peak positioned at the corresponding line parameters. Radon transform of two dimensional function  $f(x,y)$  in  $(r, \theta)$  plane is defined as

$$R(r, \theta) = R[f(x, y)]$$

$$R(r, \theta) = \iint_{-\infty}^{\infty} f(x, y) \delta(r - x \cos \theta - y \sin \theta) dx dy \quad (10)$$

Where  $\delta(\cdot)$  is Dirac function,  $r \in [-\infty, \infty]$  is perpendicular distance of a line from the origine and  $\theta \in [0, \pi]$  is the angle formed by the distance vector with x-axis as shown in figure1.

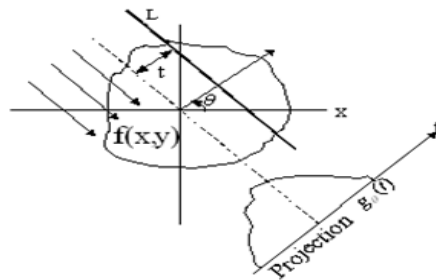


FIGURE7: Radon transform

### A. Image Fusion in Radon Space

Here we employed Radon transform for image fusion.

Algorirhm

- Register reference and sensed image.
- Compute Radon transform of both images.
- Take average of both images in radon space.
- Take inverse Radon transform.

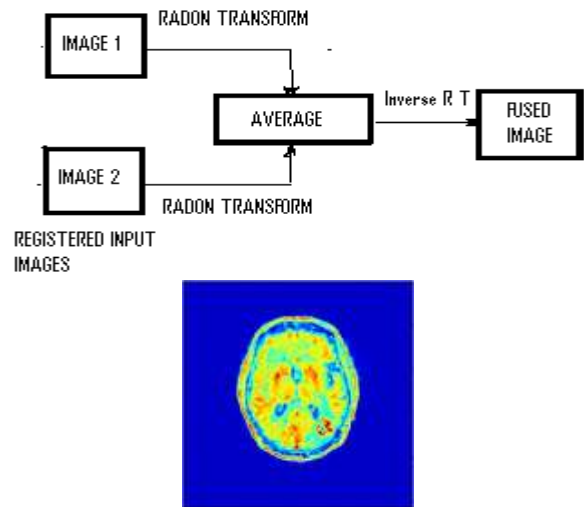


FIGURE 8: image fusion in Radon space and fused image.

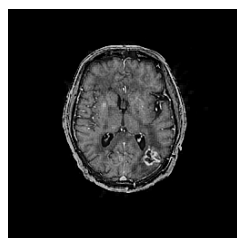
**Shift invariant discrete wavelet transform for image fusion**

The traditional DWT fusion encounters a number of shortcomings. It is well known that the DWT yields a shift variant signal representation resulting in a shift dependent fusion scheme. Fusion methods using DWT lead to unstable and flickering results. For the case of image sequences the fusion process should not be dependent on the location of object in the image and fusion output should be stable and consistent with original input sequence. SIDWT image fusion scheme overcomes this disadvantage. Considering some characteristic of the approximation wavelet coefficients of SIDWT, An approximation scale based wavelet coefficient maximum selection rule for image fusion was presented. Each stage of the SIDWT splits the input signal into the detail coefficient  $d_i(n)$ , and the approximation coefficient  $c_i(n)$  which serve as input for the next decomposition level

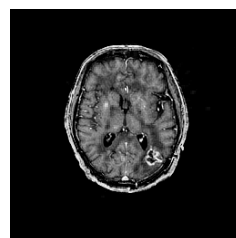
$$d_{i+1}(n) = \sum_k g(2^i \cdot k) \cdot c_i(n - k) \quad (11)$$

$$c_{i+1}(n) = \sum_k h(2^i \cdot k) \cdot c_i(n - k) \quad (12)$$

The decomposition start with  $c_0(n)=f(n)$ . The filter  $g(2^i \cdot k)$  and  $h(2^i \cdot k)$  at level  $i$  are obtained by inserting appropriate number of zeros between filter taps of the prototype filters  $g(k)$  and  $h(k)$ . The reconstruction of the input signal is performed by inverse SIDWT.



SIDWT level-2

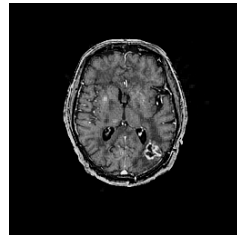


SIDWT level-4

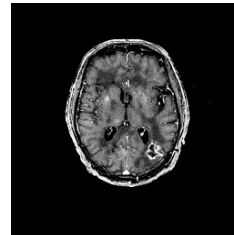
**Laplacian Pyramids image fusion**

Laplacian pyramid of an image is a set of bandpass images, in which each is a bandpass filtered copy of its predecessor. Bandpass copies can be obtained by calculating the difference between

lowpass images at successive levels of a Gaussian pyramid. In Laplacian fusion approach the Laplacian pyramids for input images are used. A strength measure is used to decide from which source what pixels contribute at each specific sample location. For example, one can use the local area sum as a measure of strength. Following figure shows Laplacian Pyramid Fusion of PET and MRI image of brain. As can be seen for both tumor and anatomy of brain can be observed in the fused image.



Laplacian fusion level-2



Laplacian fusion level-4

### Contrast Pyramid /Ratio of Low Pass Pyramid

This section introduces a hierarchical image merging scheme based on multiresolution contrast decomposition. The composite images produced by this scheme preserve those details from the input images that are most relevant to visual perception. The essential problem in merging images for visual display is "pattern conservation": important details of the component images must be preserved in the resulting fused image, while the fusion process must not introduce spurious pattern elements that could interfere with succeeding analysis.

Contrast Pyramid is similar to the ratio of Low Pass Pyramid approach. Contrast itself is defined as the ratio of the difference between luminance at a certain location in the image plane and local background luminance to the local background luminance. Luminance is defined as the quantitative measure of brightness and is the amount of visible light energy leaving a point on a surface in a given direction.

The construction of the Contrast pyramid is similar to the construction of the popular Laplacian pyramid. First a Gaussian or low-pass pyramid is constructed. This is a sequence of images in which each image is a low-pass-filtered and subsampled copy of its predecessor. Let array  $G_0$  contain the original image. This array becomes the bottom or zero level of the pyramid structure. Each node of pyramid level  $l$  ( $1 \leq l \leq N$ , where  $N$  is the index of the top level of the pyramid) is obtained as a Gaussian weighted average of the nodes at level  $l-1$  that are positioned within a  $5 \times 5$  window centered on that node.

Convolving an image with a Gaussian-like weighting function is equivalent to applying a low-pass filter to the image. Gaussian pyramid construction generates a set of low-pass-filtered copies of the input image, each with a bandwidth one octave lower than that of its predecessor. The process that generates each image in the sequence from its predecessor is

$$G_l(i,j) = \sum_{m,n=-2}^2 w(m,n) G_{l-1}(2i+m, 2j+n) \quad \text{---(13)}$$

The weighting function  $w(m,n)$  is separable:  $w(m,n) = w'(m)w'(n)$ , where  $w'(0) = a$ ,  $w'(1) = w'(-1) = 0.5$ , and  $w'(2) = w'(-2) = a/2$ . A typical value of  $a$  is 0.4. Because we are primarily interested in merging images for visual display, we demand that visually important details of the component images be preserved in the resulting composite image. It is a well-known fact that the human visual system is sensitive to local luminance contrast. If an image fusion scheme is to preserve visually important details, it must exploit this fact. We now present an image decomposition scheme that is based on local luminance contrast. This scheme computes the ratio of the low-pass images at successive levels of the Gaussian pyramid. Since these levels differ in sample

density, it is necessary to interpolate new values between the given values of the lower frequency image before it can divide the higher frequency image. Interpolation can be achieved simply by following function

$$G_{i,k}(i, j) = 4 \sum_{m,n=-2}^2 w(m, n) G_{i,k-1}\left(\frac{i+m}{2}, \frac{j+n}{2}\right) \quad \text{---(14)}$$

At every level take ratio  $R_i$  of two successive levels. Luminance contrast is defined as

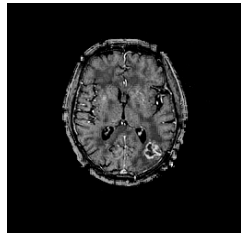
$$C = \frac{L - L_b}{L_b} = \left(\frac{L}{L_b}\right) - 1$$

- Where  $L$  - luminance at certain location of image plane
- $L_b$  - luminance of local background
- $I(i, j) = 1$  for all  $i, j$
- $C_i = R_i - 1$

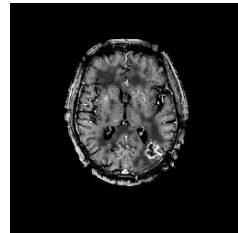
The image merging scheme can be cast into a three-step procedure. First, a ROLP pyramid is constructed for each of the source images. We assume that the different source images are in register and have the same dimensions. Second, a Ratio Of Laplace pyramid is constructed for the composite image by selecting values from corresponding nodes in the component pyramids. The actual selection rule depends on the application and may be based on individual node values or on masks or confidence estimates. For example, in the case of the fusion of two input images A and B into a single output image C and maximum absolute contrast as a selection criterion, we have, for all  $i, j$ , and  $l$

$$RC_l(i, j) = \begin{cases} RA_l(i, j), & \text{if } |RA_l(i, j) - 1| > |RB_l(i, j) - 1|, \text{ --- 15} \\ RB_l(i, j), & \text{otherwise,} \end{cases}$$

Here  $RA$  and  $RB$  represents contrast pyramids of two input images A and B and  $RC$  represents a fused image.



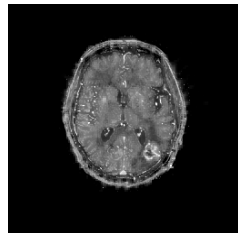
Contrast fusion level-2



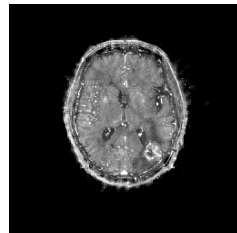
Contrast fusion level-4

**Ratio Pyramid**

Ratio of Low Pass Pyramid is another pyramid in which at every level the image is the ratio of two successive levels of the Gaussian pyramid.



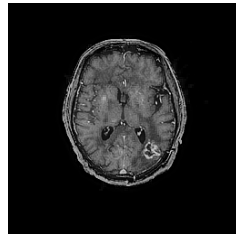
Ratio fusion level-2



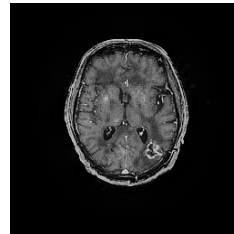
Ratio fusion level-4

### Filter-Subtract-decimate (FSD) Pyramid

FSD Pyramid technique is a more computationally efficient variation of the Gaussian Pyramid. This is similar to Laplacian fusion, the difference being in using FSD pyramid instead of Laplacian Pyramids. The only difference is in the step of obtaining the difference images in creating the pyramid. In Laplacian pyramid, the difference image  $L_k$  at level  $k$  is obtained by subtracting an image upsampled and then low-pass filtered at level  $k+1$  from the Gaussian image  $G_k$  at level  $k$ , while in FSD pyramid, this difference image is obtained directly from the Gaussian image  $G_k$  at level  $k$  subtracted by the low-pass filtered image of  $G_k$ . As a result, FSD pyramid fusion method is computationally more efficient than the Laplacian pyramid method by skipping an upsampling step.



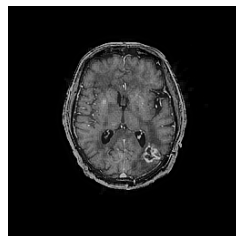
FSD Fusion level-2



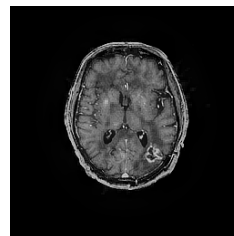
FSD Fusion level-4

### Gradient Pyramid

A gradient pyramid is obtained by applying a set of 4 directional gradient filters (horizontal, vertical and 2 diagonal) to the Gaussian pyramid at each level. At each level, these 4 directional gradient pyramids are combined together to obtain a combined gradient pyramid that is similar to a Laplacian pyramid. The gradient pyramid fusion is therefore the same as the fusion using Laplacian pyramid except replacing the Laplacian pyramid with the combined gradient.



Gradient fusion level-2



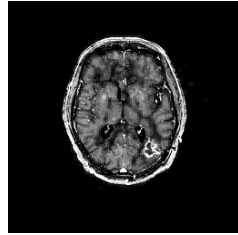
Gradient fusion level-4

### Morphological Pyramid

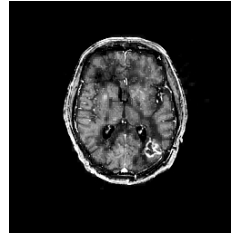
The multi-resolution techniques typically use low or bandpass filters as part of the process. These filtering operations usually alter the details of shape and the exact location of the objects in the image. This problem has been solved by using morphological filters to remove the image details without adverse effects. Morphological filters, composed of a number of elementary transformations like closing and opening transformations. The opening operator can be expressed as a composition of two other operators, erosion followed by dilation, both by the same input structural element. The main mechanism under the erosion operator is the local comparison of a shape, called structural element, Structuring element is a matrix used to define a neighborhood shape and size for morphological operations, including dilation and erosion. It consists of only 0's and 1's and can have an arbitrary shape and size. Morphological Pyramid Fusion method uses morphological pyramids instead of Laplacian or contrast pyramids.

A morphological pyramid is obtained by applying morphological filters to the Gaussian pyramid at each level and taking the difference between 2 neighboring levels. A morphological filter is usually for noise removal and image smoothing. It is similar to the effect of a low-pass filter, but it does not alter shapes and locations of objects in the image. The morphological pyramid fusion is

therefore the same as the fusion using Laplacian pyramid method except replacing the Laplacian pyramid with the morphological pyramid.



Morphological fusion level-2



Morphological fusion level-4

### 3.FUSED IMAGE EVALUATION

The purpose of multisensory fusion is to synthesize image data coming from different sensors into a single data set. Multi sensor image fusion is convenient and economical than designing the sensor with both resolution characteristics. Medical images from different sensors provide complementary information. Some applications require integration of such information. Doctors get anatomical knowledge from Magnetic Resonance Imaging (MRI) whereas physiological/functional knowledge from Photon Emission Tomography (PET). Image fusion can form a single composite image from different modality images of the same subject and provide complete information for further analysis and diagnosis. But it is necessary to align two images accurately before they fused. Before fusing images we should preserve all features in the images and should not introduce any artifacts or inconsistency which would distract the observer. Wavelet based fusion satisfies the requirement due to lots of advantages.

Image quality assessment plays an important role in medical applications. Image quality metrics are used to benchmark different image processing algorithm by comparing the objective metrics. There are two types of metrics that is subjective and objective used to evaluate image quality. In subjective metric users rate the images based on the effect of degradation and it vary from user to user whereas objective quality metrics quantify the difference in the image due to processing. The same dimension of image data is set for convenience in the fusion process and post-processing analysis. Before fusing the images they were registered. After registering, the fusion approaches- simple averaging, Principal component Analysis and wavelet based fusion at four different levels and Radon based fusion are used to create the fused images.

Assessment of image fusion performance can be first divided into two categories: one with and one without reference images. In reference-based assessment, a fused image is evaluated against the reference image which serves as a ground truth. Furthermore, fusion assessment can be classified as either qualitative or quantitative in nature. In practical applications, however, neither qualitative nor quantitative assessment alone will satisfy the needs perfectly. Given the nature of complexity of specific applications, a new assessment paradigm combing both qualitative and quantitative assessment will be most appropriate in order to achieve the best assessment result.

#### Assessment without Reference Images

In assessment without reference images, the fused images are evaluated against the original source images for similarity.

##### a. Entropy as a Quality Metric

Entropy is defined as amount of information contained in a signal. Shannon was the first person to introduce entropy to quantify the information. The entropy of the image can be evaluated as

$$H = -\sum_{i=1}^G P(i) \log_2(P(d_i)); \dots \dots \dots \langle 16 \rangle$$

Where G is the number of possible gray levels,  $P(d_i)$  is probability of occurrence of a particular gray level  $d_i$ .

Entropy can directly reflect the average information content of an image. The maximum value of entropy can be produced when each gray level of the whole range has the same frequency. If entropy of fused image is higher than parent image then it indicates that the fused image contains more information.

**b. Standard Deviation as a Quality Metric**

This metric is more efficient in the absence of noise. It measures the contrast in the fused image. An image with high contrast would have a high standard deviation.

$$\sigma = \sqrt{\sum_{i=0}^L \left( i - \bar{i} \right)^2 h_{I_f}(i), \dots \bar{i} = \sum_{i=0}^L i h_{I_f} \dots \dots \dots \langle 17 \rangle$$

Where  $h_{I_f}(i)$  is the normalized histogram of the fused image  $I_f(x,y)$  and L is number of frequency bins in histogram.

**c. Fusion Mutual Information as a Quality Metric**

It measures the degree of dependence of two images. A large measure implies better quality. If the joint histogram between  $I_1(x,y)$  and  $I_f(x,y)$  is defined as  $h_{I_1 I_f}(i,j)$  and  $I_2(x,y)$  and  $I_f(x,y)$  as  $h_{I_2 I_f}(i,j)$  then fused mutual information (FMI) is given as

$$FMI = MI_{I_1 I_f} + MI_{I_2 I_f} \dots \dots \dots \langle 18 \rangle$$

Where  $MI_{I_1 I_f} = \sum_{i=1}^M \sum_{j=1}^N h_{I_1 I_f}(i,j) \log_2 \left( \frac{h_{I_1 I_f}(i,j)}{h_{I_1}(i,j) h_{I_f}(i,j)} \right)$

$$MI_{I_2 I_f} = \sum_{i=1}^M \sum_{j=1}^N h_{I_2 I_f}(i,j) \log_2 \left( \frac{h_{I_2 I_f}(i,j)}{h_{I_2}(i,j) h_{I_f}(i,j)} \right)$$

**Assessment With Reference Image**

In reference-based assessment, a fused image is evaluated against the reference image which serves as a ground truth.

**a. Root Mean Square Error (RMSE)**

A commonly used reference-based assessment metric is the root mean square error (RMSE) which is defined as follows:

$$RMSE = \sqrt{\frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N (R(m,n) - F(m,n))^2} \dots \dots \dots \langle 19 \rangle$$



where  $R(m,n)$  and  $F(m,n)$  are reference and fused images, respectively, and  $M$  and  $N$  are image dimensions.

**b. Universal Image Quality Index ( UIQI)**

This measures how much of the salient information contained in reference image has been transformed into the fused image. The range of this metric is -1 to 1 and the best value 1 would be achieved if and only if reference and fused images are alike. The lowest value of -1 would occur when

$$I_f = 2 \mu_{I_r} - I_r$$

$$QI = \frac{4\sigma_{I_r I_f}(\mu_{I_r} + \mu_{I_f})}{(\sigma_{I_r}^2 + \sigma_{I_f}^2)(\mu_{I_r}^2 + \mu_{I_f}^2)} \quad (20)$$

Where  $\sigma_{I_r}^2$  Variance of reference image

$\sigma_{I_f}^2$  Variance of fused image

$\sigma_{I_r I_f}$  Covariance of reference and fused image

$\mu_{I_r}$  Mean of reference image

$\mu_{I_f}$  Mean of fused image

$$\mu_{I_r} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N I_r(i,j)$$

$$\mu_{I_f} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N I_f(i,j)$$

$$\sigma_{I_r}^2 = \frac{1}{MN - 1} \sum_{i=1}^M \sum_{j=1}^N (I_r(i,j) - \mu_{I_r})^2$$

$$\sigma_{I_f}^2 = \frac{1}{MN - 1} \sum_{i=1}^M \sum_{j=1}^N (I_f(i,j) - \mu_{I_f})^2$$

$$\sigma_{I_r I_f} = \frac{1}{MN - 1} \sum_{i=1}^M \sum_{j=1}^N (I_f(i,j) - \mu_{I_f})(I_r(i,j) - \mu_{I_r})$$

**c. Corelation coefficient**

The correlation coefficient measures the closeness or similarity in small size structures between the original and the fused images. It can vary between -1 and +1. Values close to +1 indicate that they are highly similar while the values close to -1 indicate that they are highly dissimilar. The ideal value is one when the reference and fused are exactly alike and it will be less than one when the dissimilarity increases.

$$CORR = \frac{2C_{rf}}{C_r + C_f} \quad (21)$$

Where

$$C_r = \sum_{i=1}^M \sum_{j=1}^N I_r(i,j)^2$$

$$C_f = \sum_{i=1}^M \sum_{j=1}^N I_f(i,j)^2$$

$$C_{rf} = \sum_{i=1}^M \sum_{j=1}^N I_f(i,j) I_r(i,j)$$

**d. Relative Mean**

The mean value of pixels in a band is the central value of the distribution of the pixels in that band. The relative shift in the mean value quantifies the changes in the histogram of the image due to processing. The relative shift in mean is defined as

$$relative\ mean = \frac{output\ mean - original\ mean}{original\ mean} \times 100\% \quad (22)$$

$$Relative\ mean = \frac{\mu_f - \mu_r}{\mu_r} \times 100\%$$

**4.RESULT**

**Data set 1 analysis**

In order to confirm the accuracy and validity of methods we have selected some sample images to carry on fusion experiment. Positron emission tomography (PET) can provide spatial information on metabolic activity in patients with cerebral glioma and functional data in patients with lesions closely related to eloquent brain areas. Because of the limited image resolution of PET, fusion with anatomic images such as MRI or CT is required to apply PET data for image guided surgery. The accuracy and clinical value of a novel image-fusion technique were evaluated for PET and MRI imaging modalities. We used the entropy of image to carry on appraisal. Table I gives the entropy of fused images and Table II shows Fusion Mutual Information.

s.n.	Method	Entropy	Mean	Std dev.
1.	Averaging	2.6333	149.6252	118.6502
2.	PCA	3.3793	23.5598	41.9296
3.	Wavelet level--2	3.3009	29.5368	57.3653
4.	Wavelet level--4	3.3653	29.9941	58.4344
5	Frequency domain	4.2764	44.6754	61.8042
6	Radon Transform	5.7492	48.5764	38.6664

7	Contrast Pyramid-level 2	2.6025	21.9640	44.1251
8	Contrast Pyramid-level 4	2.6408	23.0335	47.3534
9	FSD Pyramid-level 2	3.0847	23.4274	44.0464
10	FSD Pyramid-level 4	3.7630	24.1304	45.2177
11	Gradient Pyramid-level 2	3.0782	23.4990	44.1318
12	Gradient Pyramid-level 4	3.7470	24.2531	45.3556
13	Laplace Pyramid-level 2	2.7426	24.0742	47.7899
14	Laplace Pyramid-level 4	2.7414	26.2648	52.9269
15	Morphological Pyramid-level 2	2.9245	25.1721	52.5602
16	Morphological Pyramid-level 4	2.8756	26.1318	53.8263
17	Ratio Pyramid-level 2	2.9935	28.82	49.9349
18	Ratio Pyramid-level 4	3.0603	35.2122	60.6466
19	SIDWT with Harr level-2	2.8286	23.7187	45.8071
20	SIDWT with Harr level-4	2.7077	24.7476	50.0256

**TABLE I:** Entropy, mean and standard deviation values of fused images

s.n	Method	Combination	M.I.	FMI	Correlation Coef.	UIQI	MI	RMSE	Relative mean
1.	Simple averaging	PET-AVG <sub>f</sub>	3.0965	6.3490	0.8681	0.71153	3.4401	5.2267	0.0391%
		MRI-AVG <sub>f</sub>	3.2525						
2.	PCA	PET-PCA <sub>f</sub>	3.1002	6.2599	0.9109	0.9062	3.4355	5.2591	3.65%
		MRI-PCA <sub>f</sub>	3.1597						
3	Wavelet level-2	PET- W L2 <sub>f</sub>	3.2528	6.5465	0.8529	0.9234	3.2959	5.3566	0.81 %
		MRI- W L2 <sub>f</sub>	3.2937						
4	Wavelet level-4	PET- W L4 <sub>f</sub>	3.3392	6.7081	0.8227	0.7766	3.1840	5.4161	5.86 %
		MRI- W L4 <sub>f</sub>	3.3689						
5	Frequency domain	PET-FFT	2.7866	5.5280	0.8267	0.08397	3.2869	2.1895	1.5026 %
		MRI-FFT	2.7414						
6	Radon Transform	PET-RAD	2.5946	5.202	0.8058	0.0458	3.2597	1.9924	1.877 %

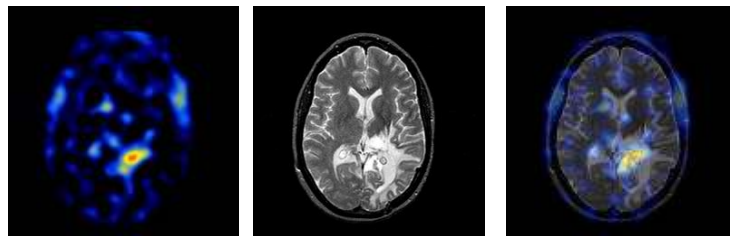
		MRI-RAD	2.607 9	5					
7	Contrast Pyramid-level 2	PET- CPL2	2.633 4	5.549 9	0.8458	0.9081	3.097 7	5.468 2	-6.70 %
		MRI- CPL2	2.916 5						
8	Contrast Pyramid-level 4	PET- CPL4	2.615 0	5.558 9	0.8368	0.9013	3.114 0	5.289 7	5.28 %
		MRI- CPL4	2.943 9						
9	FSD Pyramid-level 2	PET- FSDPL2	2.814 7	5.684 4	0.8637	0.9180	3.364 9	5.395 8	-0.50 %
		MRI- FSDPL2	2.869 7						
10	FSD Pyramid-level 4	PET- FSDPL4	2.788 2	5.650 5	0.8523	0.6930	3.302 4	5.414 8	2.50 %
		MRI- FSDPL4	2.862 3						
11	Gradient Pyramid-level 2	PET-GPL2	2.798 7	5.666	0.8651	0.9198	3.366 6	5.380 0	-0.21 %
		MRI-GPL2	2.867 3						
12	Gradient Pyramid-level 4	PET-GPL4	2.773 3	5.644 4	0.8542	0.7018	3.311 6	5.387 8	3.01 %
		MRI-GPL4	2.871 1						
13	Laplace Pyramid-level 2	PET- LPL2	2.715 7	5.566 4	0.838	0.9219	3.196 1	5.433 5	2.25 %
		MRI- LPL2	2.850 7						
14	Laplace Pyramid-level 4	PET- LPL4	2.614 9	5.417 5	0.842	0.9133	3.154 3	5.010 2	11.55 %
		MRI- LPL4	2.802 6						
15	Morphologica l Pyramid- level 2	PET-MPL2	2.798 9	5.613 1	0.7741	0.8915	3.301 4	5.516 9	6.92 %
		MRI-MPL2	2.814 2						
16	Morphologica l Pyramid- level 4	PET- MPL4	2.788 3	5.636 6	0.8032	0.8647	3.287 4	5.243 0	11.00 %
			2.848						

		MRI-MPL4	3						
17	Ratio Pyramid-level 2	PET- RPL2	2.882 9	5.716 8	0.8772	0.9160	3.326 4	3.972 2	22.42 %
		MRI- RPL2	2.833 9						
18	Ratio Pyramid-level 4	PET- RPL4	2.823 3	5.603 8	0.8620	0.8775	3.333 4	2.632 4	49.57 %
		MRI- RPL4	2.780 5						
19	SIDWT with Harr level-2	PET- SIDWTL2	2.772 1	5.631 9	0.8537	0.9200	3.265 3	5.393 9	0.72 %
		MRI- SIDWTL2	2.859 8						
20	SIDWT with Harr level-4	PET- SIDWTL4	2.629 0	5.390 9	0.8313	0.9179	3.206 8	5.397 9	5.09 %
		MRI- SIDWTL4	2.761 9						

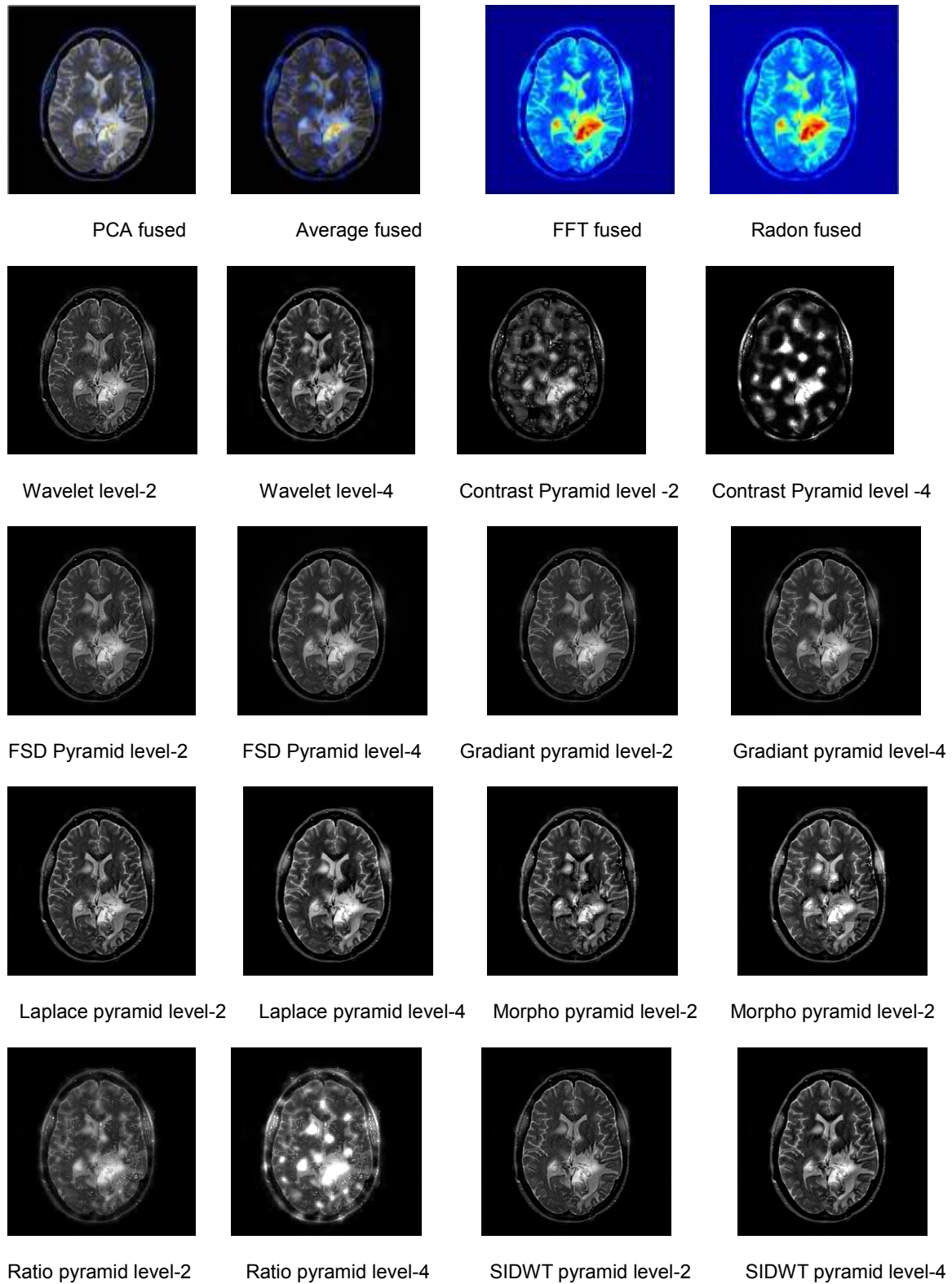
**Table II:** Fusion Mutual Information, Cross correlation, UIQI, RMSE and Relative mean

**Data set 2 analysis**

In this data set, SPECT image with tumor and MRI image of same tumor of size 256 x 256 are considered for evaluation of the fusion algorithms. It can be seen that we get Physiological/functional knowledge from Single Photon Emission Computed Tomography (SPECT) while MRI gives anatomical information. Images are fused to have both types of knowledge simultaneously. Table III gives the entropy of fused images and Table IV shows Fusion Mutual Information.



**FIGURE 11:** Sample images , SPECT (left ) and MRI (middle), reference fused image(right).



**FIGURE 12:** Images fused by different algorithms

s.n.	Method	Entropy	Mean	Std dev.
1.	Averaging	4.4555	28.4049	43.7129
2.	PCA	4.4275	35.6608	54.5568
3.	Wavelet level--2	4.9822	44.5692	55.1973
4	Wavelet level--4	5.2701	49.2849	48.1792
5	Frequency domain	5.0746	50.0557	53.9310
6	Radon Transform	4.9942	55.1852	56.3739
7	Contrast Pyramid-level 2	3.5883	19.2348	36.1107
8	Contrast Pyramid-level 4	3.4531	24.5718	39.7071
9	FSD Pyramid-level 2	4.2161	24.5201	38.5745
10	FSD Pyramid-level 4	4.5616	25.6973	40.2827
11	Gradient Pyramid-level 2	4.2117	24.5169	38.5974
12	Gradient Pyramid-level 4	4.4951	25.2924	39.8899
13	Laplace Pyramid-level 2	3.9182	25.1279	41.6786
14	Laplace Pyramid-level 4	3.7916	27.7533	47.6104
15	Morphological Pyramid-level 2	3.9905	24.5953	41.7123
16	Morphological Pyramid-level 4	4.0149	28.8634	48.7008
17	Ratio Pyramid-level 2	4.5316	29.2143	43.1095
18	Ratio Pyramid-level 4	4.7282	40.7323	60.0085
19	SIDWT with Harr level-2	4.0713	24.5718	39.7071
20	SIDWT with Harr level-4	3.7899	25.4521	43.6097

**TABLE III:** Entropy, mean and standard deviation values of fused images of dataset2

s.n.	Method	Combination	M.I.	FMI	Correlation Coef.	UIQI	MI	RMSE	Relative mean
1.	Simple averaging	SPECT-AVG <sub>f</sub>	2.7983	7.2706	0.715 1	0.4433	3.5121	6.7007	0.0419 %
		MRI-AVG <sub>f</sub>	4.4723						
2.	PCA	SPECT -PCA <sub>f</sub>	2.8556	6.7332	0.706 6	0.7886	3.5912	6.3073	-0.26 %
		MRI-PCA <sub>f</sub>	3.8776						
3	Wavelet level-2	SPECT - W L2 <sub>f</sub>	2.8851	6.708	0.690 8	0.9314	3.5286	6.1860	-.64 %
		MRI- W L2 <sub>f</sub>	3.8829						

4	Wavelet level-4	SPECT - W L4 <sub>r</sub>	2.8573	6.8999	0.671 5	0.7409	3.4754	6.8058	2.95 %
		MRI- W L4 <sub>r</sub>	3.8926						
5	Frequency domain	SPECT-FFT	2.8481	6.6918	0.693 8	0.0908	3.6194	4.747	1.024 %
		MRI-FFT	3.8437						
6	Radon Transform	SPECT-RAD	2.9007	6.709	0.686 2	0.0892	3.6081	3.7366	1.2318 %
		MRI-RAD	3.8083						
7	Contrast Pyramid-level 2	SPECT- CPL2	2.8603	6.576	0.606 6	0.8121	3.3272	0.6885	-22.2 %
		MRI- CPL2	3.7157						
8	Contrast Pyramid-level 4	SPECT- CPL4	2.7994	6.3806	0.478 6	0.6378	3.1999	6.6073	-23.30 %
		MRI- CPL4	3.5812						
9	FSD Pyramid-level 2	SPECT- FSDPL2	2.8051	6.6149	0.698 8	0.8744	3.4828	6.8316	-0.80 %
		MRI-FSDPL2	3.8098						
10	FSD Pyramid-level 4	SPECT- FSDPL4	2.8739	6.6985	0.697 2	0.6714	3.5207	6.8237	3.92 %
		MRI-FSDPL4	3.8246						
11	Gradient Pyramid-level 2	SPECT-GPL2	2.8254	6.5953	0.699 4	0.8765	3.4882	0.8309	-0.84 %
		MRI-GPL2	3.7699						
12	Gradient Pyramid-level 4	SPECT-GPL4	2.8571	6.6662	0.697 7	0.7533	3.4848	6.8504	2.30 %
		MRI-GPL4	3.8091						
13	Laplace Pyramid-level 2	SPECT- LPL2	2.7657	6.483	0.685 3	0.8268	3.4153	6.8536	1.61 %
		MRI- LPL2	3.7173						
14	Laplace Pyramid-level 4	SPECT- LPL4	2.7292	6.4592	0.681 6	0.8429	3.4098	6.6985	12.25 %
		MRI- LPL4	3.7300						
15	Morphological Pyramid-level 2	SPECT-MPL2	2.8011	6.4624	0.645 2	0.8018	3.3654	6.7190	-0.52 %
		MRI-MPL2	3.6613						
16	Morphological Pyramid-level 4	SPECT-MPL4	2.7988	6.6423	0.666 3	0.8598	3.4479	6.4851	16.74 %
		MRI-MPL4	3.8435						
17	Ratio Pyramid-level 2	SPECT- RPL2	2.9849	6.8625	0.688 2	0.8672	3.6130	6.1859	18.16 %
		MRI- RPL2	3.8776						
18	Ratio Pyramid-level 4	SPECT- RPL4	2.9843	6.9014	0.635 5	0.7024	3.5985	5.1827	64.96 %



		MRI- RPL4	3.9171						
19	SIDWT with Harr level-2	SPECT-SIDWTL2	2.8051	6.5549	0.692 2	0.8674	3.4567	6.8405	-0.60 %
		MRI-SIDWTL2	3.7498						
20	SIDWT with Harr level-4	SPECT-SIDWTL4	2.7540	6.494	0.682 7	0.8486	3.4254	6.8436	2.95 %
		MRI-SIDWTL4	3.7400						

**TABLE IV:** Fusion Mutual Information, Correlation coefficient, UIQI, RMSE, Relative mean

### 5.CONCLUSION

Image fusion is the process of image superposition using two different image types: anatomic (MRI) and functional (SPECT, positron emission tomography [PET]). This process provides the functional (SPECT antibody concentration) information in an anatomic context provided by the MRI image. It can be easy to see from the result obtained from both data sets that methods proposed in this paper have a very good effect. According to computation results, the increased entropy indicates the enhancement of information content. Approach of Radon transform provides more information as compare to other methods. But it is not clear from these results whether the enhancement of information contains more useful information or noise. From table I and III it is clear that image fusion in frequency domain and simple average contains high contrast while Radon based approaches shows a more amount of information as compared to other methods. Wavelet at level 4 shows a fair amount of contrast as well as entropy. Simple average is special domain method and produce special distortion while Fourier and Wavelet are transform domain fusions. PCA shows fair amount of entropy but contrast is not good. From table II and IV it is clear that Wavelet at level 4 shows a greater value of FMI even if Radon based approaches shows a more amount of information as compared to other methods.

Correlation coefficient was computed for different datasets. The correlation should be close to that of reference image to ensure good spectral quality. The values in the table II and IV indicate that the simple averaging and PCA method produces the best correlation result.

The histogram metrics, relative shift in mean and standard deviation, were computed. The relative shift in mean indicates the percentage by which the mean of the histogram has shifted. A positive value indicates a shift towards white and a negative value indicates a shift towards grey. The relative shift in mean values of the histogram is shown in the table II and IV for dataset 1 and dataset 2 respectively. In table IV ratio pyramid at level 2 and 4 shows 18 % and 65 % of shift respectively and in table II ratio pyramid at level 2 and 4 shows 22.42 % and 50 % shift respectively. In table IV contrast pyramid at level 2 and 4 shows -22 % and -23 % of shift respectively. This indicates a lot of distortion in the pixel values.

Table II and IV shows that fusion method based on simple averaging is least affected and wavelet, FSD pyramid, Gradient pyramid and SIDWT with Harr at level 2 are also affected less. Table II shows Root Mean Square Error value for dataset1. The values in the table indicates that the pixel values are less distorted in Radon based fusion and Frequency based method. But for dataset 2 RMSE shown in table IV indicate that pixel values are less distorted in contrast pyramid level 2 and gradient pyramid level 2.

Mutual Information between reference image and fused image also computed and it is observed that MI between reference and fused image is highest in simple averaging method for dataset 1 and ratio pyramid method for dataset 2 .are clearly visible. The simple averaging and PCA fusion

scheme leads to the best results in terms of mutual information. UIQI is highest in wavelet based method and is least in Radon based fusion.

The proposed fusion schemes were examined on two datasets .All fusion methods except Radon based fusion result in fused images of high contrast, all relevant objects of both input images are present in fused images. The SIDWT method resulted in a highly increased temporal stability of fused image sequence.

Many fusion algorithms have already been discussed to evaluate quality of fused image; however, these algorithms have not been assessed in terms of their visual performance. The visual performance-based assessment methodology is equally important. We evaluated the quality of a fused image by comparing its visual differences with the source images and which require no knowledge of the ground truth. It is observed that Wavelet at level-4 gives complete information with a better degree of contrast but at the cost of time. PCA and wavelet based fusion have been traditionally used for image fusion, but they have their own shortcomings. For PCA most information is gained if all calibrated and visually good quality multi-spectral input bands are used. PCA is general purpose approach and not application oriented. Wavelet based fusion can deal with images of different spectral and spatial resolutions. However, this method cannot handle cases where data is scattered or when input images differ greatly in either their spectral or spatial resolution. Figure 12 shows fused images of MRI image and SPECT image of brain with tumor. Brain tumor is clearly visible in SPECT image but not in MRI image . Figure shows that although single-photon emission computerized tomography (SPECT) provides unique functional information it is limited by poor anatomic detail. Non-specific localization of the antibody makes image interpretation difficult, since contrast improvement offered by SPECT imaging increases both the tumor specific and non-specific background uptake. Anatomy of brain can be observed in MRI image , not in SPECT image, but fused image shows anatomy of brain as well as tumor. The essential problem in merging images for visual display is “ pattern conservation”. Important details of the component image must be preserved in the resulting fused image; while the merging process must not introduce spurious pattern elements that could interfere with subsequent analysis. Laplacian decomposition scheme is based on local luminance differences. However, the human visual system is sensitive only to local luminance contrast. In the example of SPECT and MRI image if observers interest is in tumor then contrast pyramid is a better choice but if observers interest is in anatomy as well as tumor both, then SIDWT, yields good results. Researchers concluded that nonspecific antibody concentration in the blood pool or other normal anatomic structures is more easily recognized with MRI-SPECT or MRI-PET fusion due to the spatial overlay on structure. They also indicated that MRI-SPECT fusion was also useful in identifying tumor masses that were undetected on the MRI alone due to adjacent underfilled small bowels. The advantages of image fusion over visual comparison of multimodality are: (a) the fusion technique is useful to correct for variability in orientation, position and dimension; (b) it allows precise anatomic-physiologic correlation; and (c) it permits regional quantitation. Tumor volumes defined on the basis of antibody concentration can be generated from the SPECT images. The distribution of antibody uptake in specific anatomic regions or non-specific background regions can also be evaluated. Three-dimensional volumes of tumors of sufficient size can be generated and viewed from different spatial perspectives. Another important use of the registration-fusion technique is in follow-up studies where the primary purpose is to identify recurrence of the tumors or the spreading of metastatic disease. The major consequences of the fusion process is that the sites of suspected activity can be revisited post-surgery to determine the extent of recovery or recurrence of disease.

Future work includes more algorithms that combine aspect of pixel level and feature level fusion and more quality metrics to evaluate fused images.

## 6. REFERENCES

1. Brown Gottesfeld L., "Survey of Image Registration techniques", ACM Computing Surveys, 24, 4, and 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. Rafael C. Gonzalez, R. E. Woods, Stevan L. Eddins, *Digital Image Processing* (Pearson Education, 2003).
5. 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.
6. 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 2:14-17 Dec. 2003 ,2: 976 - 981
7. 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, 2: 758-760
8. Du-Ming Tsai, Ron-Hwa Yang "An eigenvalue-based similarity measure and its application in defect detection" Image and Vision Computing , 23(12): 1094-1101, Nov 2005,
9. P.A. Van den Elsen, Evert-Jan D. Pol et al. "Medical Image Matching- a Review with Classification." IEEE Engineering Medicine and Biology, March , pp 26-38, 1993.
10. H. Li, B. S. Manjunath and S. K. Mitra "Multisensor Image Fusion Using the Wavelet Transform", Graphical Models and Image Processing. 57 (3): 235-245, 1995.
11. Zhao Zong-gui "An Introduction to Data Fusion Method." First press. 28<sup>th</sup> Institute of Electricity Ministry, 1998.
12. JIA Yong-hong, Li de-ren, SUN Jia-bing "Multidimensional Remote Sensing Imagery Data Fusion" Remote Sensing Technology and Application 2005, 15 (1) : 41-44.
13. David A Y "Image Merging and Data Fusion by Means of the Discrete Two Dimensional Wavelet Transform" J. Opt. Soc. Am. A, 1995 , 12 (9) : 1834-1841.
14. S. Mallat, "A Wavelet Tour of Signal Processing" Academic Press, Second Edition, 1998.
15. Svensson, J. R., M. O. Ulfarsson & J. A. Benediktsson, "Cluster Based Feature Extraction and Data Fusion in the Wavelet Domain" In: Pro. IEEE International Geoscience and Remote Sensing Symposium, pp. 867-869.
16. Garzelli, A "Possibilities and Limitations of the Use of Wavelets in Image Fusion." In: Pro. IEEE International Geoscience and Remote Sensing Symposium, 2002.
17. Lau Wai Leung, Bruce King and Vijay Vohora, "Comparison of Image Fusion Techniques using Entropy and INI", In: Pro. 22<sup>nd</sup> Asian Conference on Remote Sensing, 5-9 Nov 2001.

18. Chavez, P.S., Sides, S. C. Anderson, J.A., "Comparison of Three Different Methods to Merge Multiresolution and Multispectral Data: Landsat TM and SPOT Panchromatic", Photogrammetric Engineering and Remote Sensing, 57,295-303.
  19. 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.
  20. Lau Wai Leung, Bruce King and Vijay Vohora, "Comparison of Image Fusion Techniques using Entropy and IN", In: Pro. 22<sup>nd</sup> Asian Conference on Remote Sensing, 5-9 Nov 2001.
  21. Rockinger, O., "Image Sequence Fusion Using a Shift Invariant Wavelet Transform," Proceedings of the International conference on Image Processing, 1997.
  22. Ramac, L. C., Uner, M. K., Varshney, P. K., "Morphological filters and wavelet based image fusion for concealed weapon detection," Proceedings of SPIE, 3376, 1998.
  23. Nunez, J., "Multiresolution-based image fusion with additive wavelet decomposition," IEEE Transactions on Geoscience and Remote Sensing, 37(3), 1999.
  24. Alexander Toet, van Ruyven, J.J. & Valetton, J.M., "Merging Thermal and Visual Images by a Contrast Pyramid", Optical Engineering, 28(7), pp. 789-792.
  25. Vivek Maik, Jeongho Shin and Joonki Paik, "Pattern Selective Image Fusion for Focus Image Reconstruction", Caip 2005, LNCS 3691, pp 677-684, 2005.
  26. Wen Doua, Yunhao Chen, "An Improved Image Fusion Method with High Spectral Fidelity", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B7. Beijing 2008.
  27. Andrew P. Bradley, "Shift-invariance in the Discrete Wavelet Transform" Proc. VIIIth Digital Image Computing: Techniques and Applications, Sun C., Talbot H., Ourselin S. and Adriaansen T. (Eds.), 10-12 Dec. 2003, Sydney.
  28. Milad Ghantous, Soumik Ghosh and Magdy Bayoumi, "A Gradient-based Hybrid Image Fusion Scheme using Object Extraction", 2008 IEEE, 978-1-4244-1764-3/08.
  29. Toet, J. Van Ruvan, and J. Valetton, "Merging thermal and visual images by a contrast pyramid," Optical Engineering, 28, 1989.
  30. Alexander Toet, "Multiscale contrast enhancement with applications to image fusion", Optical Engineering 31(5), 1026-1031 (May 1992).
  31. P. Burt, E. Adelson, "Laplacian pyramid as a compact image code," IEEE Transactions on Communications, 31(4), 1983.
  32. Xydeas, C., and Petrovic, V., "Objective Pixel-level Image Fusion Performance Measure," Sensor Fusion: Architectures, Algorithms, and Applications IV, SPIE 4051:89-98, 2000.
-

## A Novel Image Retrieval System Using an Effective Region Based Shape Representation Technique

**Santhosh P Mathew**

*Professor/Computer Science & Engineering  
Saintgits College of Engineering  
Kottayam, Kerala, India PIN 686 532*

mathewsantosh@yahoo.com

**Philip Samuel**

*Reader & Head/Information Technology  
Cochin University of Science & Technology  
Cochin, Kerala, India PIN 682 022*

philipsamu@yahoo.com

---

### Abstract

With recent improvements in methods for the acquisition and rendering of shapes, the need for retrieval of shapes from large repositories of shapes has gained prominence. A variety of methods have been proposed that enable the efficient querying of shape repositories for a desired shape or image. Many of these methods use a sample shape as a query and attempt to retrieve shapes from the database that have a similar shape. This paper introduces a novel and efficient shape matching approach for the automatic identification of real world objects. The identification process is applied on isolated objects and requires the segmentation of the image into separate objects, followed by the extraction of representative shape signatures and the similarity estimation of pairs of objects considering the information extracted from the segmentation process and shape signature. We compute a 1D shape signature function from a region shape and use it for region shape representation and retrieval through similarity estimation. The proposed region shape feature is much more efficient to compute than other region shape techniques invariant to image transformation.

**Keywords:** Shape representation, Image retrieval, shapes signature, enhancement, segmentation

---

## 1. INTRODUCTION

An image retrieval system is a computer system for browsing, searching and retrieving images from large repositories of digital images. The increase in social web applications and the semantic web have inspired the development of several web-based image annotation tools[1]. The processing semantics of the images is an open research area, as a support for a series of complex operations in the field of processing images and videos, such as recognizing shapes and objects, understanding video and detect potential risk events in video surveillance, etc. For semantic retrieval based on image content, it is important to have existing metadata such as an annotation for images. For each domain of application, where image retrieval is desired, semantic relationship needs to be established between the objects (the structure of the working object) present in the images[2,3]. Image search is a specialized data search used to find images. To search for images, a user may provide query terms such as keyword, image file/link, or click on some image, and the system will return images "similar" to the query[4]. Here we propose a system in which query is an image and the output of it is images related to that query i.e. images that contains input image.

In this paper, we propose a novel retrieval system using contour based method for region based shape representation and retrieval [7]. Conventional shape methods use grid sampling to acquire shape information. The shape representation derived this way is usually not translation, rotation and scaling invariant. Extra normalization is therefore required. Goshtasby proposes the use of shape matrix which is derived from a circular raster sampling technique [8]. The idea is similar to normal raster sampling. However, rather than overlaying the normal square grid on a shape image, a circular sampling of concentric circles and radial lines is overlaid at the centre of mass of the shape. The binary value of the shape is sampled at the intersections of the circles and radial lines. The shape matrix is formed so that the circles correspond to the matrix columns and the radial lines correspond to the matrix rows. The result matrix representation is invariant to translation, rotation, and scaling.

Specifically, we compute a one dimensional signature function from a region shape and use it for region shape representation and retrieval [9]. We call it circular raster sampling signature or CRSS for short. We show that the proposed technique is robust and outperforms the widely used shape methods in literature.

## 2. RELATED WORKS

Conventional information retrieval is based solely on text, and these approaches to textual information retrieval have been transplanted into image retrieval in a variety of ways, including the representation of an image as a vector of feature values. Wilkins et al propose a technique that uses text based IR methods for indexing MPEG-7 visual features to perform rapid subset selection within large image collections [10]. Barrios et al present an image retrieval system based on a combined search of text and content. The idea is to use the text present in title, description, and tags of the images for improving the results obtained with a standard content-based search [13]. To provide text descriptions or annotations, two approaches can be applied. The first approach acquires descriptions/annotations manually by human annotators. The second approach is to automatically annotate images using machine-learning techniques that learn the correlation between image features and textual words from the examples of annotated images [14, 15].

However, “a picture is worth a thousand words.” Image contents are much more versatile compared with text, and the amount of visual data is already enormous and still expanding very rapidly. Hoping to cope with these special characteristics of visual data, content-based image retrieval methods have been introduced. It has been widely recognized that the family of image retrieval techniques should become an integration of both low-level visual features, addressing the more detailed perceptual aspects, and high-level semantic features underlying the more general conceptual aspects of visual data. Neither of these two types of features is sufficient to retrieve or manage visual data in an effective or efficient way. Although efforts have been devoted to combining these two aspects of visual data, the gap between them is still a huge barrier in front of researchers. Intuitive and heuristic approaches do not provide us with satisfactory performance. Therefore, there is an urgent need of finding and managing the latent correlation between low-level features and high-level concepts. How to bridge this gap between visual features and semantic features has been a major challenge in this research field.

## 3. APPROACH OVERVIEW

The system proposed is an efficient content based image retrieval system. This basically has three sections: Signature Generation, Image Storage and Image Retrieval. First, signatures are generated for the enhanced and segmented image, by applying the circular grid. Then, we store the set of signatures of the basic image patterns in our reference table. Finally, we input a query image to the system and find its corresponding set of signatures and search with the stored set of signature in reference table. If we get signatures which are matching, then the system outputs all

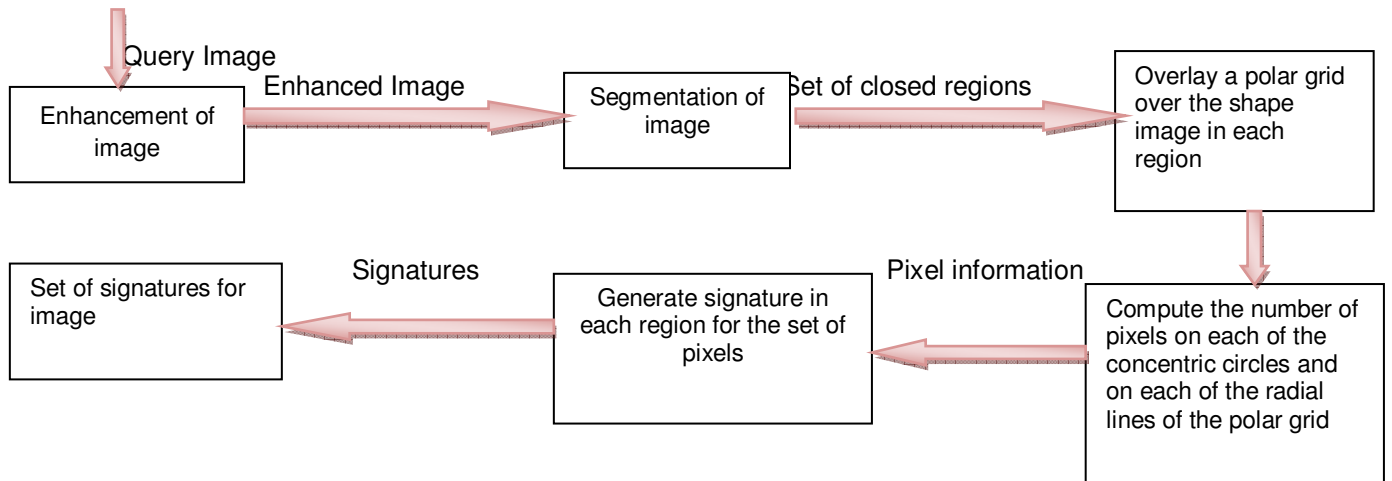
the images that are mapped with matching reference signature. And if not, it will be added to the reference table.

**Steps for Signature Generation based on CRSS Algorithm**

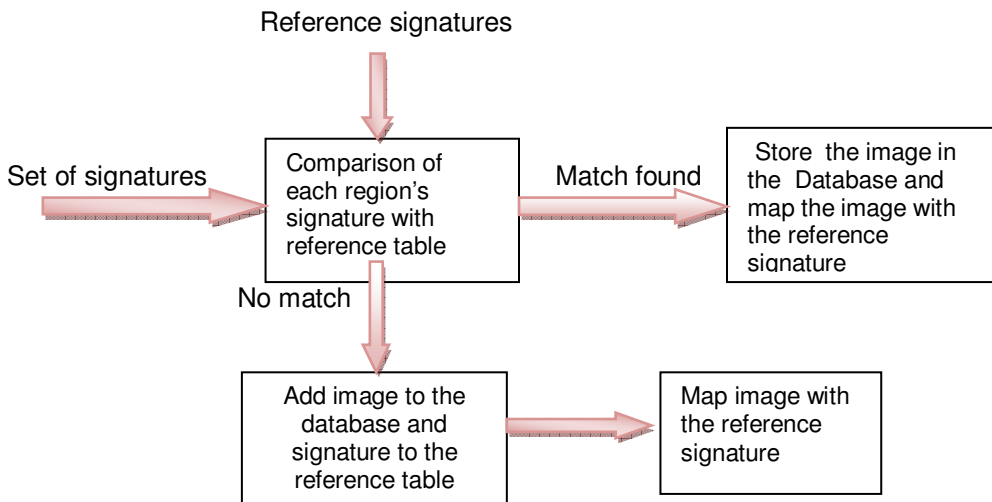
- Step 1. Input image
- Step 2. Enhancement of image to reduce noise and increase its image quality.
- Step 3. Segment the image into different closed regions.
- Step 4. For each image do following
  - 4.a Determine the centroid of image.
  - 4.b Overlay polar grid over shape image to compute number of shape pixels on intersection of concentric circles and radial lines for each region
  - 4.c Collect pixel information – pixel position of these pixels
  - 4.d. Generate signature from the set of shape pixels.

Figure 1 describes the block diagram of the proposed retrieval system.

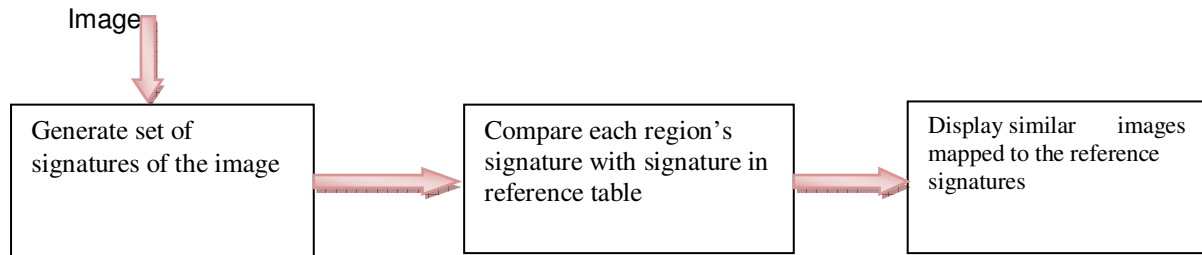
**Signature Generation**



**Storing Image to the Database**



## Image Retrieval



## 4. IMAGE ENHANCEMENT

Image enhancement processes consist of a collection of techniques that seek to improve the visual appearance of an image or to convert the image to a form better suited for analysis by a human or machine. Meanwhile, the term image enhancement is mean as the improvement of an image appearance by increasing dominance of some features or by decreasing ambiguity between different regions of the image. Contrast stretching is the image enhancement technique that is commonly used. To date, contrast stretching process plays an important role in enhancing the quality and contrast of images. There are 4 steps involved in applying image enhancement process. a) The first step is image capturing b) Then, save the images under .bmp extension. c) The third step is to select picture with 3 different types which is normal image, bright image and dark image. Three images are selected for each different type. d) The last step is applying the identified Partial Contrast stretching technique to the selected images.

Partial contrast stretching is an auto scaling method. It is a linear mapping function that is usually used to increase the contrast level and brightness level of the image. This technique will be based on the original brightness and contrast level of the images to do the adjustment. The mapping function is as follows [11,12]:

$$P_k = ((\max - \min) / (f_{\max} - f_{\min})) * (q_k - f_{\min}) + \min$$

Where,

$P_k$  : color level of the output pixel  $q_k$  : color level of the input pixel

$f_{\max}$  : maximum color level values in the input image  $f_{\min}$  : minimum color level values in the input image  $\max$  &  $\min$  : desired maximum and minimum color levels that determines color range of the output image, respectively

Before the mapping process start, the system will find the range of where the majority of the input pixels converge for each colour space. Since the input images are the RGB model, so it is necessary to find the range for the red, blue and green intensities. After that, the average will be calculated for these upper and lower colour values of the range of three colour space by using the following formula [12]:

$$\maxTH = (\maxRed + \maxBlue + \maxGreen) / 3$$

$$\minTH = (\minRed + \minBlue + \minGreen) / 3$$

$\maxRed$ ,  $\maxBlue$  and  $\maxGreen$  are the maximum colour level for each red, blue and green colour palettes, respectively.  $\minRed$ ,  $\minBlue$  and  $\minGreen$  are the minimum value for each colour palette, respectively.  $\maxTH$  and  $\minTH$  are the average number of these maximum and minimum colour levels for each colour space. The  $\maxTH$  and  $\minTH$  will be used as the desired colour ranges for all the three colour palettes. The purpose of the three colour palette to have the same threshold value is to avoid the colour level to be placed out side of a valid colour level. After



that, the mapping process will start [12]. The function in Equation 4 will be used for the pixels transformation, which is based on the concept of the linear mapping function in Equation .

$$out(x, y) = \begin{cases} \frac{in(x,y)}{minTH} * NminTH & \text{for } in(x,y) > minTH \\ \left[ \frac{NMaxTH - NMinTH}{maxTH - minTH} * (in(x,y) - fmin) \right] + \min & \text{for } minTH < in(x,y) < maxTH \\ \frac{in(x,y)}{maxTH} * NmaxTH & \text{for } in(x,y) < maxTH \end{cases}$$

where,  $in(x,y)$  : colour level for the input pixel  
 $out(x,y)$  : colour level for the output pixel  
 $minTH$  : lower threshold value  
 $maxTH$  : upper threshold value  
 $NminTH$  : new lower stretching value  
 $NmaxTH$  : new upper stretching value

The pixel within the range of  $minTH$  and  $maxTH$  will be stretched to the desire range of  $NmaxTH$  to  $NminTH$ , whereas the remaining pixels will experience compression. By this stretching and compressing processes, the pixels of the image can be mapped to a wider range and brighter intensities; as a result the contrast and the brightness level of the raw images are increased. Figure 1 illustrates the compression and stretching processes for partial contrast method. The value of 80 and 200 were used as an example of lower and upper threshold value while 20 to 230 as the desired range of the colour level for the output image. The original range of the input image will be stretched to the range from 20 to 230. The colour level below 80 will be compressed to the range of 0 to 20 and the colour level more than 200 will be compressed to the range of 230 to 255.

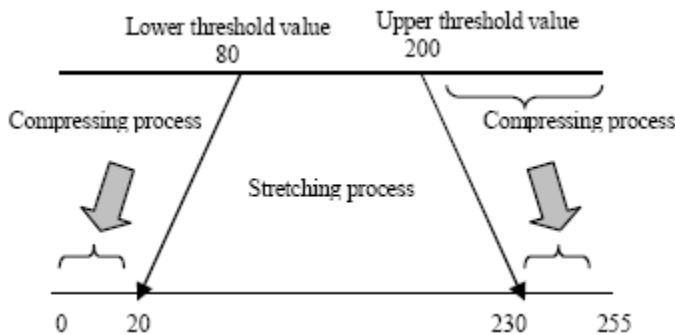


Figure 2 Partial Contrast Stretching Method

## 5. IMAGE SEGMENTATON

Segmentation, i.e. the partitioning of image data into related sections or regions, is a key first step in a number of approaches to data analysis and compression. In image analysis, the group of image data points contained in each region provides a statistical sampling of image data values for more reliable labeling based on image feature values[16]. In addition, the region shape can be analyzed as an additional clue for the appropriate labeling of the region. In data compression, the regions form a basis for compact representation of the image data. The quality of the prerequisite image segmentation is a key factor in determining the level of performance of most of these

image analysis and data compression approaches. Region growing approaches to segmentation are preferred here because region growing exploits spatial information and Guarantees the formation of closed, connected regions [5].

A less commonly used approach to region growing image segmentation is the Hierarchical Stepwise Optimization algorithm of Beaulieu and Goldberg [6]. HSWO is best defined iteratively: Start with an image and a segmentation of that image into  $N$  regions in which (i) every picture element (pixel) is in a region, (ii) and each region is connected, (*i.e.* composed of contiguous image pixels). Then compare all spatially adjacent regions with each other (*e.g.*, compute a vector norm between the region means of the spatially adjacent regions). Merge the most similar pair of spatially adjacent regions. Continue to compare spatially adjacent regions and merge the most similar pair of spatially adjacent regions until either a specified number of regions are reached or the dissimilarity between the most similar pair of spatially adjacent regions reaches a specified threshold.

The initial partition may assign each image pixel to a separate region. Any other initial partition may be used, such as an over-segmented result from region growing based on the classic definition given above (*i.e.*, classic region growing segmentation with a low threshold value). Here we make use of the recursive region growing approach utilized by the hierarchical image segmentation (RHSEG) algorithm described by James Tilton[5]. It is identical to that employed by Beaulieu and Goldberg's HSWO algorithm except that HSEG optionally alternates spectral clustering iterations with region growing iterations. In the spectral clustering iterations, non-adjacent regions are merged. In fact, in their paper on HSWO, Beaulieu and Goldberg [6] provide the theoretical basis for the HSEG algorithm in their theoretical analysis of HSWO. They show that the HSWO algorithm produces the globally optimal segmentation result if each iteration is statistically independent. Even though each iteration will generally not be statistically independent for natural images, the HSWO approach still produces excellent results. Beaulieu and Goldberg also point out that the sequence of partitions generated by this iterative approach reflect the hierarchical structure of the imagery data: the partitions obtained in the early iterations preserve the small details and objects in the image, while the partitions obtained in the latter iterations preserve only the most important components of the image. They further note that these hierarchical partitions may carry information that help in identifying the objects in the imagery data.

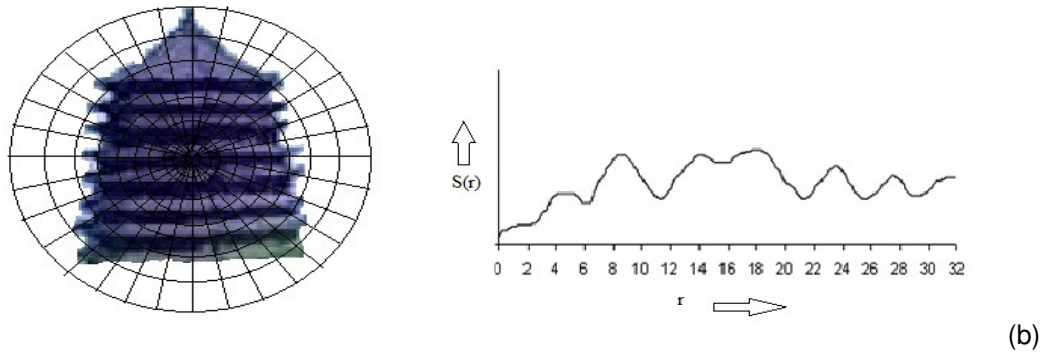


**FIGURE 3:** Image Segmentation using RHSEG

## 6. CIRCULAR RASTER SAMPLING SIGNATURE

In this section, we introduce the basic idea of the shape representation technique named as *circular raster sampling signature (CRSS)*, and describe its implementation details. The shape matrix technique captures local shape information well, however, it suffers from noise and high

dimension matching drawbacks. The area-ratio technique is robust to noise, however, it does not capture local shape information, as the result, it is not accurate[17]. In this context, we propose a polar raster sampling signature technique which computes a signature function of the sampled points. Specifically, we overlay a polar raster grid over the shape image, and compute the number of shape pixels on each of the concentric circles and on each of the diameters of the polar sampling grid. Then the shape pixels are filtered to avoid noise. This is done by calculating the intensity of the pixel and comparing with a threshold value. The number of pixels is a function of the radius and the angle, the function is called circular raster sampling signature, or CRSS for short. In the following, we describe the proposed technique in details.



**FIGURE 4:** (a) polar sampling of a shape (b)  $S(r)$  component of the CRSS of (a)

### 6.1 Implementation

To implement PRS, a shape image needs to be converted from Cartesian space to polar space. Given an shape image in Cartesian space:  $I = \{f(x, y); 0 \leq x \leq M, 0 \leq y \leq N\}$ , it can be represented in polar space as  $I_p = \{f(r, \theta); 0 \leq r \leq R, 0 \leq \theta \leq 2\pi\}$ , where  $R$  is the maximum radius of the shape. The origin of the polar space, set to be the centroid of the shape, is initialised to 0.

**Step 1:** Find the centroid  $(x_c, y_c)$  of the region, which is given by

$$x_c = \frac{1}{M} \sum_{x=0}^{N-1} x, \quad y_c = \frac{1}{N} \sum_{y=0}^{M-1} y$$

**Step 2:** Calculation of pixels position  $(x, y)$  of each shape pixel on each of the concentric circles and radial lines.( for each region)

**Step 3:** Find  $(r, \theta)$  is given by:

$$r = \sqrt{(x - x_c)^2 + (y - y_c)^2}, \quad \theta = \arctan \frac{y - y_c}{x - x_c}$$

**Step 4:** The PRS signature function consists of two components:  $S(r)$  and  $S(\theta)$ , which are given by

$$S(r) = \sum_x \sum_y f(x, y) \delta \left[ r - \sqrt{(x - x_c)^2 + (y - y_c)^2} \right]$$

$$S(\theta) = \sum_x \sum_y f(x, y) \delta \left[ \theta - \arctan \frac{y - y_c}{x - x_c} \right]$$

Where  $0 \leq r \leq R$ ,  $0 \leq \theta \leq \pi$ .  $f(x, y)$  is the binary value of image pixel at  $(x, y)$ ,  $\delta(x)$  is the Dirac delta function:

$$\delta(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } x \neq 0 \end{cases}$$

In our case,  $r$  is quantized by 32, and  $\theta$  is quantized by 30. An example shape and its first component of CRSS signature are shown in Figure 4.

## 7. CONCLUSION & FUTURE WORK

In this paper, we have proposed an efficient and robust content based image retrieval system using the region based shape representation technique named *circular raster shape signature (CRSS)*. The novel approach uses contour shape method to compute a concentric circular sampling signature from a region shape which may not have a continuous boundary. It overcomes several drawbacks in existing techniques, like noise sensitivity, loss of local information and computational complexity. The proposed shape descriptor has better performance compared with well known shape representation techniques and will be useful in retrieval of similar image from large repositories of images.

The semantic interpretation of the image and the spacing of the concentric circles in the circular grid can further be studied to improve the storage and retrieval performance of the CRSS based image retrieval System.

## 8. REFERENCES

1. K. S. Fu and J. K. Mui, "A survey on image segmentation," *Pattern Recognition*, Vol 13, pp. 3-16, 1981.
2. J. Eakins, M. Graham: Content-based Image Retrieval. Technical Report, University of Northumbria at Newcastle, 1999.
3. F. Long, H. J. Zhang and D. D. Feng: Fundamentals of Content-based Image Retrieval. In D. Feng Eds, *Multimedia Information Retrieval and Management—Technological Fundamentals and Applications*, Springer, 2003.
4. Y. Rui, T. S. Huang, and S.-F. Chang: Image Retrieval: Current Techniques, Promising Directions, and Open Issues. *Journal of Visual Communication and Image Representation*, 10(4):39-62, 1999.
5. James C. Tilton, "Method for recursive hierarchical segmentation by region Growing and spectral clustering with a natural convergence criterion," *Disclosure of Invention and New Technology: NASA Case No. GSC 14,328-1*.
6. Jean-Marie Beaulieu and Morris Goldberg, "Hierarchy in picture segmentation: A stepwise optimization approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 2, pp. 150-163, February 1989.
7. D. S. Zhang and G. Lu: Review of Shape Representation and Description Techniques. *Pattern Recognition*, 37(1):1-19, 2004.
8. A. Goshtasby: Description and Discrimination of Planar Shapes Using Shape Matrices *IEEE Trans on Pattern Recognition and Machine intelligence - Vol 6 November 1985*

- 9]. Dengsheng Zhang and Melissa Chen Yi Lim. An Efficient and Robust Technique for Region Based Shape Representation 6th IEEE/ACIS International Conference on Computer and Information Science (ICIS 2007)
10. Wilkins, Peter and Ferguson, Paul and Smeaton, Alan F. and Gurrin, Cathal (2005) *Text based approaches for content-based image retrieval on large image collections*. In: EWIMT 2005 - 2nd European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies, 30 Nov -1 Dec 2005, London, UK.
11. R.W.Jr. Weeks,(1996). *Fundamental of Electronic Image Processing*. Bellingham: SPIE Press.
12. Nor Hazlyna Harun ,N.R.Mokhtar,M.Y. Mashor , H.Adilah ,R.Adollah,Nazahah Mustafa, N.F.Mohd Nasir , H.Roseline, 'Color image enhancement techniques based on partial contrast and contrast stretching for acute leukaemia images', ICPE-2008.
13. Juan Manuel Barrios, Diego Díaz-Espinoza, Benjamin Bustos, "Text-Based and Content-Based Image Retrieval on Flickr: DEMO," *sisap*, pp.156-157, 2009 Second International Workshop on Similarity Search and Applications, 2009
14. Blei, D. and Jordan, M. Modeling annotated data. In *Proceedings of 26th International Conference on Research and Development in Information Retrieval (SIGIR)*. 2003.
15. Jeon, J., Lavrenko, V., and Manmatha, R. Automatic Image Annotation and Retrieval using Cross-Media Relevance Models. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*. 2003.
16. Raman Maini, Himanshu Aggarwal "Study and Comparison of Various Image Edge Detection Techniques" *International Journal of Image Processing (IJIP)*, 3(1):1-11, 2009
17. Chandra Sekhar Panda, Srikanta Patnaik, "Filtering Corrupted Image and Edge Detection in Restored Grayscale Image Using Derivative Filters", *International Journal of Image Processing (IJIP)*,3(3):105-119

# CALL FOR PAPERS

**Journal:** International Journal of Image Processing (IJIP)

**Volume:** 4 **Issue:** 6

**ISSN:** 1985-2304

**URL:** <http://www.cscjournals.org/csc/description.php?JCode=IJIP>

## About IJIP

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.

## IJIP 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
- 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

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

## **IMPORTANT DATES**

**Volume:** 4

**Issue:** 6

**Paper Submission:** November 31, 2010

**Author Notification:** January 01, 2011

**Issue Publication:** January /February 2011

## CALL FOR EDITORS/REVIEWERS

CSC Journals is in process of appointing Editorial Board Members for ***International Journal of Image Processing (IJIP)***. CSC Journals would like to invite interested candidates to join **IJIP** network of professionals/researchers for the positions of Editor-in-Chief, Associate Editor-in-Chief, Editorial Board Members and Reviewers.

The invitation encourages interested professionals to contribute into CSC research network by joining as a part of editorial board members and reviewers for scientific peer-reviewed journals. All journals use an online, electronic submission process. The Editor is responsible for the timely and substantive output of the journal, including the solicitation of manuscripts, supervision of the peer review process and the final selection of articles for publication. Responsibilities also include implementing the journal's editorial policies, maintaining high professional standards for published content, ensuring the integrity of the journal, guiding manuscripts through the review process, overseeing revisions, and planning special issues along with the editorial team.

A complete list of journals can be found at <http://www.cscjournals.org/csc/byjournal.php>. Interested candidates may apply for the following positions through <http://www.cscjournals.org/csc/login.php>.

*Please remember that it is through the effort of volunteers such as yourself that CSC Journals continues to grow and flourish. Your help with reviewing the issues written by prospective authors would be very much appreciated.*

Feel free to contact us at [coordinator@cscjournals.org](mailto:coordinator@cscjournals.org) if you have any queries.



## **Contact Information**

### **Computer Science Journals Sdn Bhd**

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

Phone: +603 6207 1607  
          +603 2782 6991  
Fax:     +603 6207 1697

### **BRANCH OFFICE 1**

Suite 5.04 Level 5, 365 Little Collins Street,  
MELBOURNE 3000, Victoria, AUSTRALIA


Fax: +613 8677 1132

### **BRANCH OFFICE 2**

Office no. 8, Saad Arcad, DHA Main Bulevard  
Lahore, PAKISTAN

### **EMAIL SUPPORT**

Head CSC Press: [coordinator@cscjournals.org](mailto:coordinator@cscjournals.org)  
CSC Press: [cscpress@cscjournals.org](mailto:cscpress@cscjournals.org)  
Info: [info@cscjournals.org](mailto:info@cscjournals.org)



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