

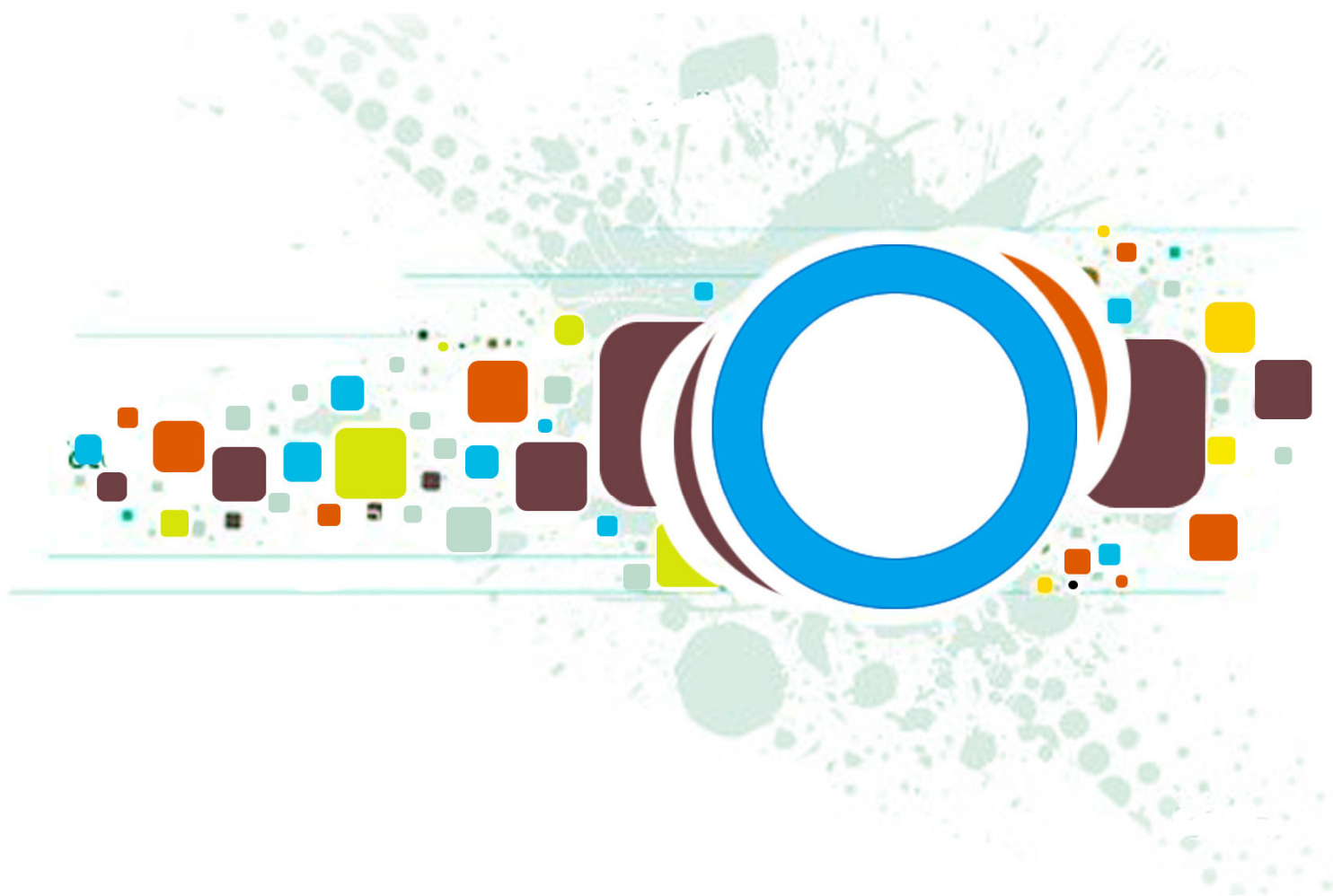
Volume 7 • Issue 2 • April 2013

Editor-in-Chief
Professor Hu, Yu-Chen

INTERNATIONAL JOURNAL OF
IMAGE PROCESSING (IJIP)

ISSN : 1985-2304

Publication Frequency: 6 Issues Per Year



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

INTERNATIONAL JOURNAL OF IMAGE PROCESSING (IJIP)

VOLUME 7, ISSUE 2, 2013

**EDITED BY
DR. NABEEL TAHIR**

ISSN (Online): 1985-2304

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

IJIP Journal is a part of CSC Publishers

Computer Science Journals

<http://www.cscjournals.org>

INTERNATIONAL JOURNAL OF IMAGE PROCESSING (IJIP)

Book: Volume 7, Issue 2, April 2013

Publishing Date: 30-04-2013

ISSN (Online): 1985-2304

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

IJIP Journal is a part of CSC Publishers

<http://www.cscjournals.org>

© IJIP Journal

Published in Malaysia

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

CSC Publishers, 2013

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 second issue of volume seven of IJIP. The Journal is published bi-monthly, with papers being peer reviewed to high international standards. IJIP emphasizes on efficient and effective image technologies, and provides a central for a deeper understanding in the discipline by encouraging the quantitative comparison and performance evaluation of the emerging components of image processing. IJIP comprehensively cover the system, processing and application aspects of image processing. Some of the important topics are architecture of imaging and vision systems, chemical and spectral sensitization, coding and transmission, generation and display, image processing: coding analysis and recognition, photopolymers, visual inspection etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 7, 2013, IJIP appears in more focused issues. Besides normal publications, IJIP intends to organize special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

IJIP gives 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. Iftekharruddin
University of Memphis
United States of America

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

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

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

Professor Yen-Wei Chen
Ritsumeikan University
Japan

Associate Professor Tao Gao
Tianjin University
China

Dr Choi, Hyung Il
Soongsil University
South Korea

EDITORIAL BOARD MEMBERS (EBMs)

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

Dr Ghassan Adnan Hamid Al-Kindi
Sohar University
Oman

Dr Cho Siu Yeung David

Nanyang Technological University
Singapore

Dr. E. Sreenivasa Reddy
Vasireddy Venkatadri Institute of Technology
India

Dr Khalid Mohamed Hosny
Zagazig University
Egypt

Dr Chin-Feng Lee
Chaoyang University of Technology
Taiwan

Professor Santhosh.P.Mathew
Mahatma Gandhi University
India

Dr Hong (Vicky) Zhao
Univ. of Alberta
Canada

Professor Yongping Zhang
Ningbo University of Technology
China

Assistant Professor Humaira Nisar
University Tunku Abdul Rahman
Malaysia

Dr M.Munir Ahamed Rabbani
Qassim University
India

Dr Yanhui Guo
University of Michigan
United States of America

Associate Professor András Hajdu
University of Debrecen
Hungary

Assistant Professor Ahmed Ayoub
Shaqra University
Egypt

Dr Irwan Prasetya Gunawan
Bakrie University
Indonesia

Assistant Professor Concetto Spampinato
University of Catania
Italy

Associate Professor João M.F. Rodrigues

University of the Algarve
Portugal

Dr Anthony Amankwah

University of Witswatersrand
South Africa

Dr Chuan Qin

University of Shanghai for Science and Technology
China

Associate Professor Vania Vieira Estrela

Fluminense Federal University (Universidade Federal Fluminense-UFF)
Brazil

Dr Zayde Alcicek

firat university
Turkey

Dr Irwan Prasetya Gunawan

Bakrie University
Indonesia

TABLE OF CONTENTS

Volume 7, Issue 2, April 2013

Pages

- 109 - 123 An Assessment of Image Matching Algorithms in Depth Estimation
Ashraf Anwar, Ibrahim El Rube
- 124 - 131 Performance Evaluation of Object Tracking Technique Based on Position Vectors
V Purandhar Reddy, K Thirumala Reddy, Y.B.Dawood, K.M.YaagnaTheja, C.Munesh, M.V.L.Sraavani
- 132 - 139 Comparative Analysis of Partial Occlusion Using Face Recognition Techniques
Nallammal.N, V.Radha
- 140 - 148 Image Fusion Quality Assessment of High Resolution Satellite Imagery based on an Object Level Strategy
Farhad Samadzadegan, Farzaneh Dadras Javan
- 149 - 162 ISEF Based Identification of RCT/Filling in Dental Caries of Decayed Tooth
Ami Jigar Solanki, K. R. Jain, N. P. Desai
- 163 - 170 Recognition of Facial Expressions using Local Binary Patterns of Important Facial Parts
Ramchand Hablani, Narendra Chadhari, Sanjay Tanwani
- 171 - 182 A Comprehensive Survey on Human Facial Expression Detection
Archana Verma, Lokesh Kumar Sharma
- 183 - 190 Texture features from Chaos Game Representation Images of Genomes
Vrinda V Nair, Nisha N. S., Vidya S., Y. S. Thushana

- 191 - 202 Diagnosis of Burn Images using Template Matching, k-Nearest Neighbor and Artificial Neural Network
Malini Suvarna, Sivakumar, Kamal Kumar, U. C. Niranjana
- 203 - 208 Target Detection by Fuzzy Gustafson-Kessel Algorithm
Mousumi Gupta
- 209 - 218 HABIT: Handwritten Analysis based Individualistic Traits Prediction
Abdul Rahiman, Diana Varghese, Manoj Kumar G
- 219 - 226 Brain Tumor Extraction from T1- Weighted MRI using Co-clustering and Level Set Methods
S.Satheesh, K.V.S.V.R Prasad, K.Jitender Reddy

An Assessment of Image Matching Algorithms in Depth Estimation

Ashraf Anwar

*College of Computers and Information Technology
Taif University, Taif, Saudi Arabia*

ashraaf@tu.edu.sa

Ibrahim El Rube

*College of Computers and Information Technology
Taif University, Taif, Saudi Arabia*

ielrube@yahoo.com

Abstract

Computer vision is often used with mobile robot for feature tracking, landmark sensing, and obstacle detection. Almost all high-end robotics systems are now equipped with pairs of cameras arranged to provide depth perception. In stereo vision application, the disparity between the stereo images allows depth estimation within a scene. Detecting conjugate pair in stereo images is a challenging problem known as the correspondence problem. The goal of this research is to assess the performance of SIFT, MSER, and SURF, the well known matching algorithms, in solving the correspondence problem and then in estimating the depth within the scene. The results of each algorithm are evaluated and presented. The conclusion and recommendations for future works, lead towards the improvement of these powerful algorithms to achieve a higher level of efficiency within the scope of their performance.

Keywords: Stereo vision, Image Matching Algorithms, SIFT, SURF, MSER, Correspondence.

1. INTRODUCTION

Stereo vision systems are used to determine depth from two images taken at the same time but from slightly different viewpoints using two cameras. The main aim is to calculate disparity which indicates the difference in locating corresponding pixels in two images. From the disparity map, we can easily calculate the correspondence of objects in 3Dspace which is known as depth map. The known algorithms for stereo matching can be classified in two basic categories: Feature-based algorithms and area based algorithms [2-14]. The algorithms of both categories often use special methods to improve the matching reliability.

Matas et al [14] find maximally stable extremely regions (MSER) correspondences between image elements from two images with different viewpoints. This method of extracting a comprehensive number of corresponding image elements contributes to the wide-baseline matching, and it has led to better stereo matching and object recognition algorithms. David. G and Lowe [15] proposed a scale invariant feature transform (SIFT) detector and descriptor, which detects a set of local feature vectors through scale space extremes and describe this feature using 3D histogram of gradient and orientation. Also Hess [16] introduced an open source SIFT library. Herbert Bay, et. al, [17] proposed SURF (Speeded-Up Robust Features) as a fast and robust algorithm for local, similarity invariant image representation and comparison. SURF selects interest points of an image from the salient features of its linear scale-space, and then builds local features based on the image gradient distribution. An open source SURF library is introduced by Evans, C. [18].

In this paper, we propose to use one of the well known image matching algorithms (SIFT, MSER, or SURF) in estimating the distance between the SVS surveyor robot, shown in figure 1, and the in front obstacles. Therefore, an evaluation and assessment of the performance of the three

image matching algorithms is conducted to determine the best applicable algorithm for the SVS surveyor robot.

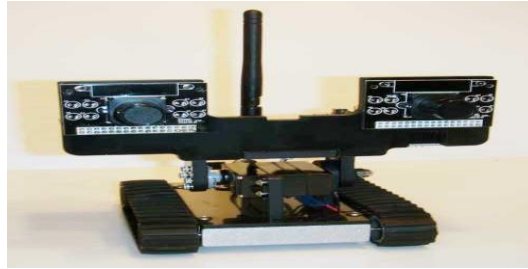


FIGURE 1: SVS Surveyor Robot.

This paper is organized as follows: section 2 presents the overview on the SVS surveyor robot. Stereo vision concept illustrated in section 3. Section 4 is dedicated to image matching algorithms. Section 5 is reserved to the methodology. Simulation results are presented in section 6. Finally section 7 concludes this paper.

2. SVS SURVEYOR ROBOT

The SVS surveyor robot [1] is designed for research, education, and exploration, Surveyor's internet-controlled robot. The robot is usually equipped with two digital video cameras with resolution from 160x128 to 1280x1024 pixels, two laser pointers, and WLAN 802.11b/g networking on a quad-motor tracked mobile robotic base.

Operating as a remotely-controlled webcam or a self-navigating autonomous robot, the robot can run onboard interpreted C programs or user-modified firmware, or be remotely managed from a Windows, Mac OS/X or Linux base station with Python or Java-based console software.

2.1 Stereo Vision System Specifications

The two SRV-1 Blackfin camera modules are separated by 10.75 cm (4.25"). Each camera module includes:

- 500MHz Analog Devices Blackfin BF537 Processor (1000 integer MIPS), 32MB SDRAM, 4MB SPI Flash, JTAG, external 32-pin i/o header w/ 2 UARTS, 4 timers (PWM/PPM), SPI, I2C, 16 GPIO
- Omnivision OV9655 1.3 megapixel sensor with AA format header and interchangeable lens - M12 P0.5 format - 3.6mm f2.0 (90-deg FOV) or optional 2.2mm f2.5 (120-deg FOV)

3. STERO VISION CONCEPT

3.1 Basics

The geometric basis key problem in stereo vision is to find corresponding points in stereo images. Corresponding points are the projections of a single 3D point in the different image spaces. The difference in the position of the corresponding points in their respective images is called disparity (see figure 2). Two cameras: Left and Right, Optical centers: OL and OR. Virtual image plane is projection of actual image plane through optical centre. Baseline, b , is the separation between the optical centers. Scene Point, P , imaged at PL and PR . Disparity, $d = PR - PL$.

Disparity is the amount by which the two images of P are displaced relative to each other
Depth, $Z = bf/p*d$

Where p : pixel width

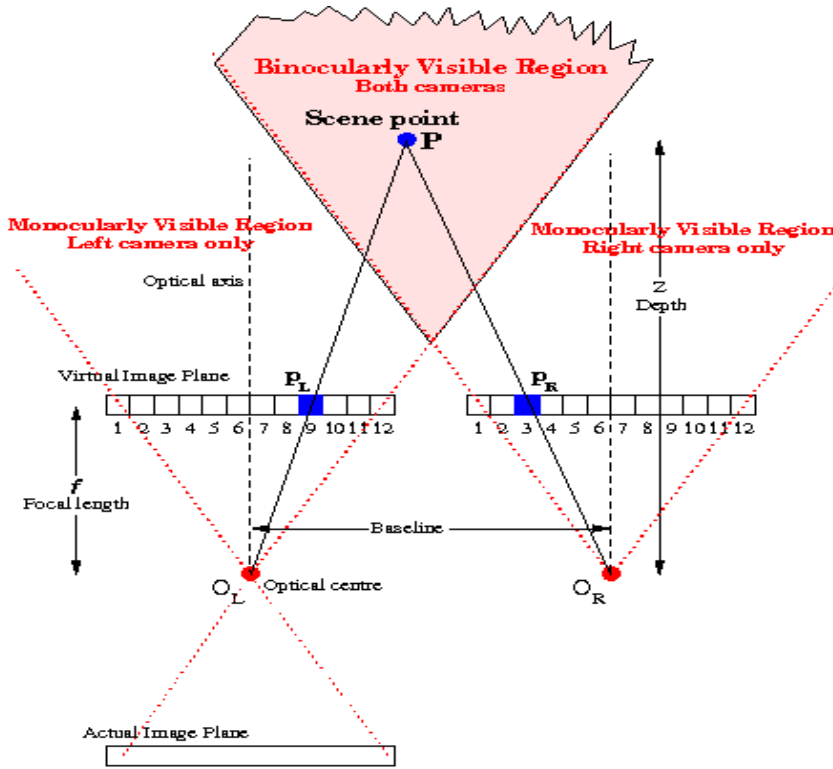


FIGURE 2: Stereo Vision Basics.

In addition to providing the function that maps pair of corresponding images points onto scene points, a camera model can be used to constraint the search for corresponding image point to one dimension. Any point in the 3D world space together with the centers of projection of two cameras systems, defines an epipolar plane. The intersection of such a plane with an image plane is called an epipolar line (see figure 3). Every point of a given epipolar line must correspond to a single point on the corresponding epipolar line. The search for a match of a point in the first image therefore is reduced to a one-dimensional neighborhood in the second image plane.

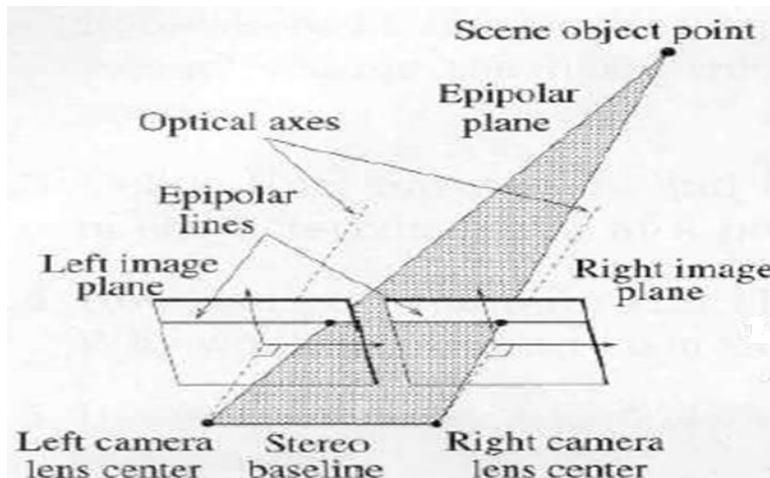


FIGURE 3: Epipolar Lines and Epipolar Planes.

3.2 Correspondence problem

There are two issues, how to select candidate matches? and how to determine the goodness of a match? Two main classes of correspondence (matching) algorithm: First, Correlation-based, attempt to establish a correspondence by matching image intensities, usually over a window of pixels in each image. Second Feature-based, attempt to establish a correspondence by matching sparse sets of image features, usually edges. Disparity map is sparse, and number of points is related to the number of image features identified. Feature-based methods, suitable when good features can be extracted from the scene, faster than correlation-based methods, provide sparse disparity maps, suitable for applications like visual navigation, and relatively insensitive to illumination changes.

4. IMAGE MATCHING ALGORITHMS

4.1 The Scale Invariant Feature Transform (SIFT) Algorithm

The SIFT algorithm operates in four major stages [16, 20] to detect and describe local features, or keypoints, in an image:

- Scale-space extrema detection. The SIFT algorithm begins by identifying the locations of candidate keypoints as the local maxima and minima of a difference-of-Gaussian pyramid that approximates the second order derivatives of the image's scale space.
- Keypoint localization and filtering. After candidate keypoints are identified, their locations in scale space are interpolated to sub-unit accuracy, and interpolated keypoints with low contrast or a high edge response computed based on the ratio of principal curvatures are rejected due to potential instability.
- Orientation assignment. The keypoints that survive filtering are assigned one or more canonical orientations based on the dominant directions of the local scale-space gradients. After orientation assignment, each keypoint's descriptor can be computed relative to the keypoint's location, scale, and orientation to provide invariance to these transformations.
- Descriptor computation. Finally, a descriptor is computed for each keypoint by partitioning the scale-space region around the keypoint into a grid, computing a histogram of local gradient directions within each grid square and concatenating those histograms into a vector. To provide invariance to illumination change, each descriptor vector is normalized to unit length, threshold to reduce the influence of large gradient values, and then renormalized.

For image matching and recognition, SIFT features are first extracted from a set of reference images and stored in a database. A new image is matched by individually comparing each feature from the new image to this previous database and finding candidate matching features based on Euclidean distance of their feature vectors..

4.2 The Maximally Stable Extremely Regions (MSER)

It is a feature detector; Like the SIFT detector, the MSER algorithm extracts from an image a number of co-variant regions, called MSERs. An MSER is a *stable* connected component of some level sets of the image. Optionally, elliptical frames are attached to the MSERs by fitting ellipses to the regions. Because the regions are defined exclusively by the intensity function in the region and the outer border, this leads to many key characteristics of the regions which make them useful. Over a large range of thresholds, the local linearization is stable in certain regions, and have the properties listed below.

- Invariance to affine transformation of image intensities

- Covariance to adjacency preserving (continuous) transformation $T : D \rightarrow D$ on the image domain
 - Stability: only regions whose support is nearly the same over a range of thresholds is selected.
 - Multi-scale detection without any smoothing involved, both fine and large structure is detected.
- Note however that detection of MSERs in a scale pyramid improves repeatability, and number of correspondences across scale changes.

This technique was proposed by Matas et al. [14] to find correspondences between image elements from two images with different viewpoints. This method of extracting a comprehensive number of corresponding image elements contributes to the wide-baseline matching, and it has led to better stereo matching and object recognition algorithms.

4.3 Speeded Up Robust Features (SURF)

It is a robust local feature detector, first presented by Herbert Bay et al. [17, 20], it can be used in computer vision tasks like object recognition or 3D reconstruction. SURF is based on sums of 2D Haar wavelet responses and makes an efficient use of integral images.

The steps of features detection as follows:

- Interest points are selected at distinctive locations in the image, such as corners, blobs, and T-junctions. The most valuable property of an interest point detector is its repeatability, i.e. whether it reliably finds the same interest points under different viewing conditions.
- Next, the neighborhood of every interest point is represented by a feature vector. This descriptor has to be distinctive and, at the same time, robust to noise, detection errors, and geometric and photometric deformations.
- Finally, the descriptor vectors are matched between different images. The matching is often based on a distance between the vectors, e.g. the Mahalanobis or Euclidean distance. The dimension of the descriptor has a direct impact on the time this takes, and a lower number of dimensions is therefore desirable.

5. METHODOLOGY

5.1 Data source

Two SRV-1 Blackfin camera modules, illustrated in section 2

5.2 Camera Calibration

The result of camera calibration using the Camera Calibration Toolbox for Matlab [19] is obtained in Tables 1 and 2 for left and right cameras of the SVS stereo system mentioned in section 2, respectively.

Focal Length	fc_left = [390.97269 371.48472] ± [90.30192 86.25323]
Principal point:	cc_left = [176.99127 -0.14410] ± [0.00000 0.00000]
Skew	alpha_c_left = [0.00000] ± [0.00000] => angle of pixel axes = 90.00000 ± 0.00000 degrees
Distortion	kc_left = [1.03881 -2.69365 -0.01420 0.03846 0.00000] ± [1.38449 5.07636 0.12276 0.02306 0.00000]

TABLE 1: Intrinsic Parameters of Left Camera.

Focal Length	fc_right = [490.50860 470.70292] ± [94.29747 97.23895]
Principal point:	cc_right = [159.50000 119.50000] ± [0.00000 0.00000]
Skew	alpha_c_right = [0.00000] ± [0.00000] => angle of pixel axes = 90.00000 ± 0.00000 degrees
Distortion	kc_right = [-0.76164 9.78187 0.18255 0.00095 0.00000] ± [1.30013 11.26048 0.09762 0.01903 0.00000]

TABLE 2: Intrinsic Parameters of Right Camera.

5.3 Test Algorithm

The framework of the proposed algorithm is summarized in the following steps:

- Step 1: Read stereo image pair.
- Step 2: Compute interest points for each image using SURF/SIFT/MSER algorithm.
- Step 3: Find point correspondences between the stereo image pair.
- Step 4: Remove outliers using geometric constraint.
- Step 5: Remove further outliers using Epipolar constraint.
- Step 6: Rectify images such that the corresponding points will appear on the same rows.
- Step 7: Obtain the disparity map and calculate the depth.

5.4 Processing Steps.

In order to assess the performance of the three image matching algorithms, SIFT, MSER, and SURF, we applied the proposed algorithm on a set of images captured by stereo vision system of SVS surveyor robot. The image-pairs are captured at different distances, 50cm, 100cm, 150cm, 200cm, 250cm, and 300cm, as shown in figures 4 to 9 (a-l, a-r; ;b-l, b-r; c-l, c-r; d-l, d-r; e-l, e-r; f-l, f-r).

5.5 Performance Evaluation.

The effectiveness of the algorithm is calculated using the following formula:

$$E\% = \frac{\# \text{ correct matches}}{\min (\# \text{left features}, \# \text{rightFeatures})} 100 \quad (1)$$

The depth accuracy is calculated according to the following formula:

$$\text{Depth Acc \%} = \frac{\text{Real Depth} - \text{Min. Depth}}{\text{Real Depth}} 100 \quad (2)$$



FIGURE 4: Image pair at 50cm from stereo camera



FIGURE 5: Image pair at 100cm from stereo camera



c-l



c-r

Figure 6: Image Pair at 150cm From Stereo Camera.



d-l



d-r

FIGURE 7: Image Pair at 200cm From Stereo Camera.



FIGURE 8: Image Pair at 250cm From Stereo Camera.

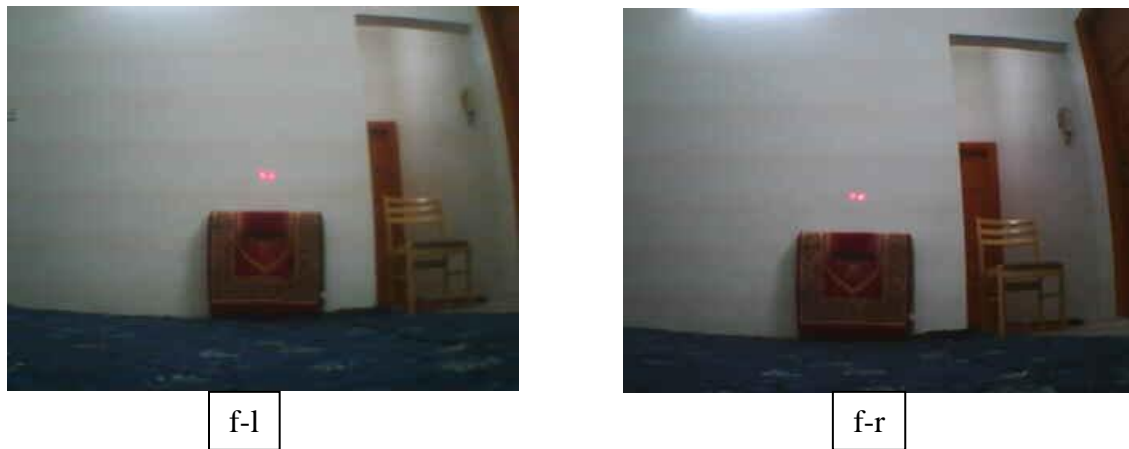


Figure 9: Image Pair at 300cm From Stereo Camera.

6. SIMULATION RESULTS

6.1 Platform

The simulation is performed using matlab software (R2012a). The computer processor is Intel® core TM, i5, M430, 2.27 GHz. The matching algorithms, SURF, SIFT, and MSER are tested individually by every image-pair, illustrated in figure 4 to figure 9 according to the steps of processing explained in section 5.4. The image results of every step are shown in figure 10. The performance results of SIFT, MSER, and SURF, are tabulated in Table 3, Table 4, and Table 5 respectively. The bar plots for features detected, effectiveness, and depth accuracy, for the three matching algorithms, are illustrated in figure 11, figure 12, and figure 13 respectively.

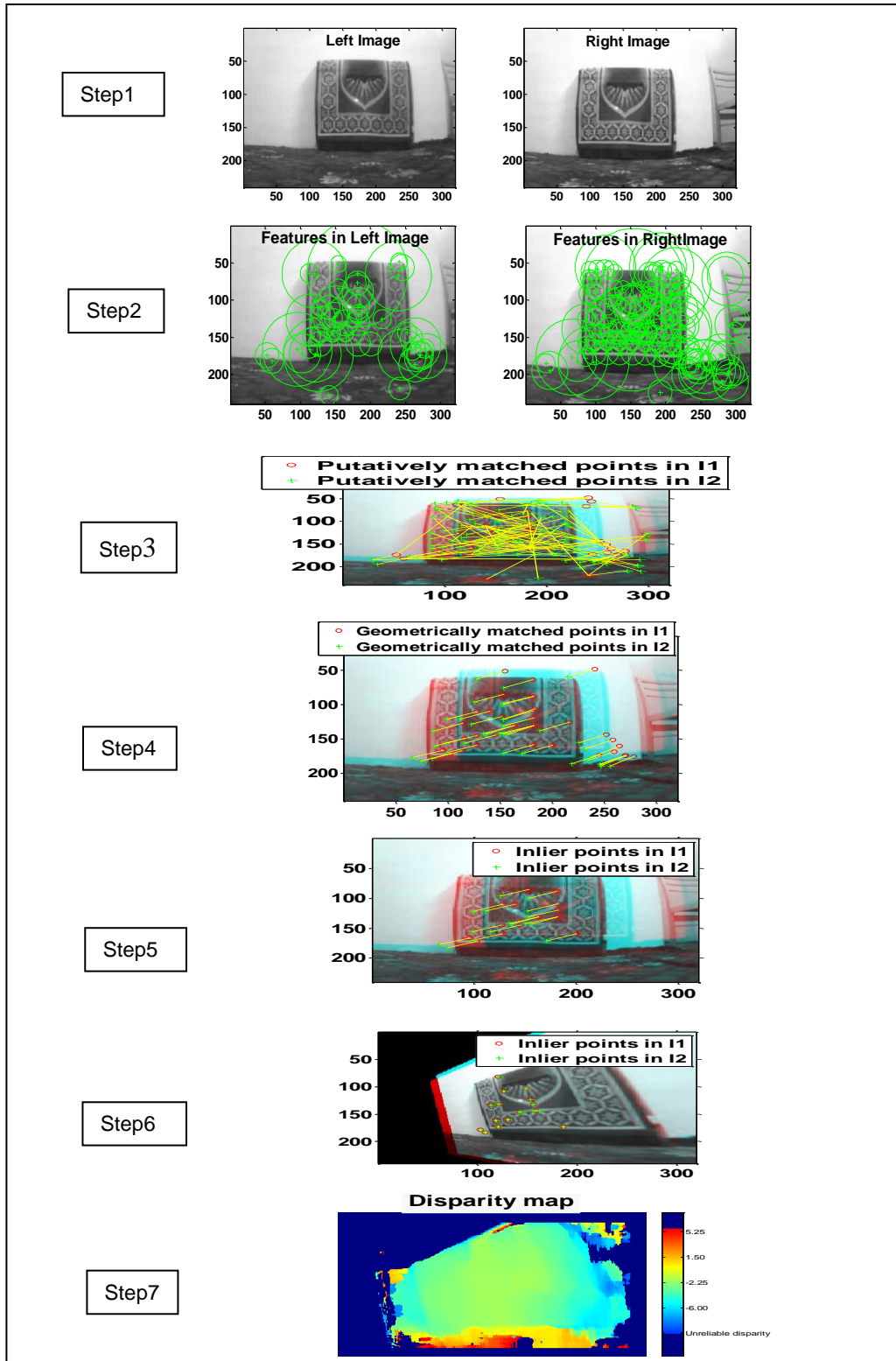


FIGURE10: Image Results.

Image pair	Real depth (cm)	Features		Match features	#Correct matches	E%	Min. depth(cm)	Depth acc%
		Left	Right					
a	50	713	753	178	68	9.5	63.1	73.8
b	100	638	646	198	106	16.6	116.8	83.2
c	150	664	513	136	54	10.5	165.3	89.8
d	200	640	644	101	28	4.3	172.8	86.4
e	250	594	609	62	15	2.5	156.7	62.7
f	300	582	429	85	17	3.9	263.7	87.9

TABLE 3: SIFT Matching Results.

Image pair	Real depth (cm)	Features		Match features	# Correct matches	E%	Min. depth(cm)	Depth acc%
		Left	Right					
a	50	186	220	342	18	9.6	51.1	97.8
b	100	164	193	277	42	25.6	93.5	93.5
c	150	65	127	157	20	30.7	133.3	88.8
d	200	70	74	118	14	20	168.2	84.0
e	250	34	34	58	10	29.4	130	52.0
f	300	25	41	53	12	48	135	45.0

TABLE 4: MSER Matching Results.

Image pair	Real depth (cm)	Features		Match features	# Correct matches	E%	Min. depth (cm)	Depth acc%
		Left	Right					
a	50	299	335	482	27	9.0	50.2	99.6
b	100	153	267	314	26	16.9	94.1	94.1
c	150	48	128	139	17	35.4	140.8	93.8
d	200	27	32	45	12	44.4	171.8	85.9
e	250	21	29	37	9	42.8	213.9	85.5
f	300	22	27	34	9	40.9	234.1	78.0

TABLE 5: SURF Matching Results.

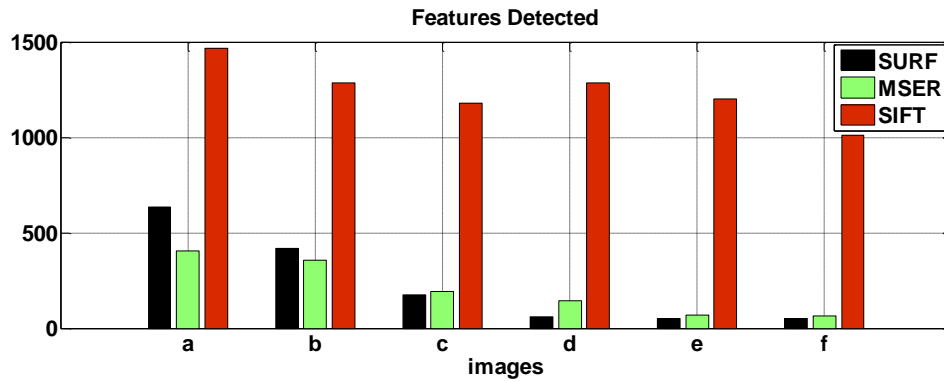


Figure 11: Bar Plot of the Detected Features.

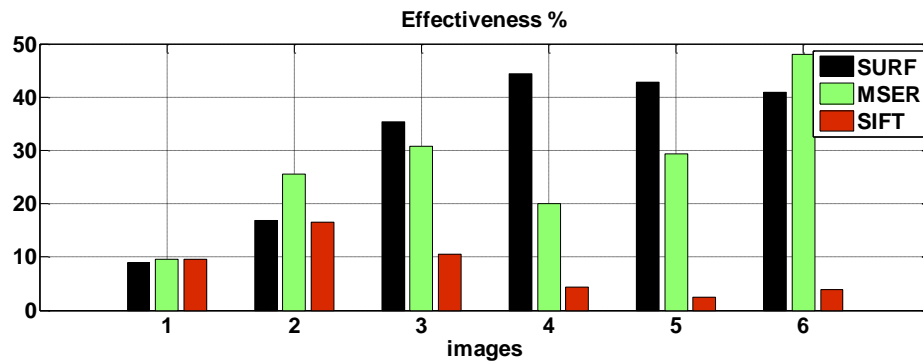


FIGURE 12: Bar Plot of the Algorithms Effectiveness.

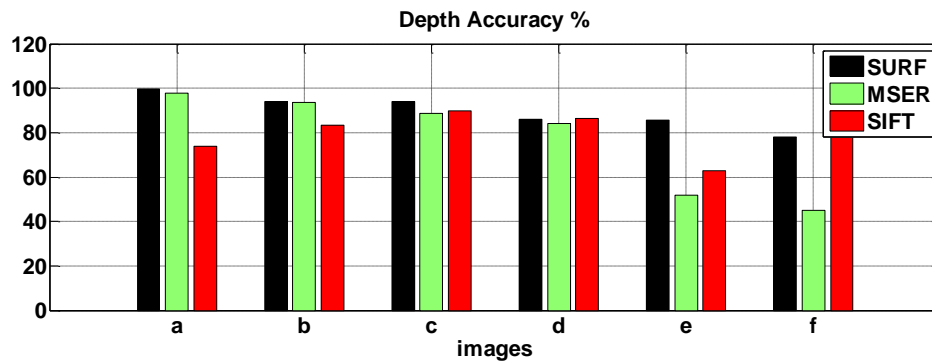


FIGURE13: Bar Plot of the Depth Accuracy.

6.2 Analysis of the Results

After reviewing the results obtained, we have the following notes:

- a) The amount of features detected by SIFT, MSER and SURF is dependent on the depth between the stereo camera and the object.
- b) In case of SURF, we decrease the metric threshold to 500 instead of 1000 (default value), when the distance greater than 200 cm in order to get appropriate features.

- c) In case of MSER, we decrease the threshold delta to min value (0.8) instead of 2 (default value), when the distance greater than 200 cm in order to get appropriate features.
- d) The amount of features is not a measure of success by itself but the “quality” of these features
- e) The amount of features detected is proportional to the amount of matches.
- f) Although the SIFT detect more matches, but the SURF gives the best result in estimating the depth in all images. We can deduce that, the amount of matches detected is not a good indication of the performance of the algorithm.
- g) Matches detected by SURF, although fewer, are more robust than those detected by SIFT and MSER.

6.3 Comparative Evaluation

To evaluate the performance of the our test algorithm in the depth estimation, we get also the maximum depth in case of using SURF as a matching algorithm (which gives the best results), and then we get the average depth between the minimum depth and the maximum depth. The results are given in Table 6.

Real depth (cm)	Estimated depth (cm)		Mean depth (cm)	Average absolute error (cm)
	Min.	Max.		
50	50.2	53.4	51.8	1.8
100	94.1	110.1	102.1	2.1
150	140.8	168.5	154.65	3.6
200	171.8	235.8	203.8	3.8
250	213.9	302.7	258.3	8.3
300	234.1	391.1	312.6	12.6

TABLE 6: Estimated Depth Results.

The comparison between the results obtained by Young [21] method in distance estimation and our test algorithm is illustrated in Table 7.

Real depth (cm)	Young [21] Method		Our test algorithm	
	Average absolute error(cm)	% Average absolute error	Average absolute error(cm)	% Average absolute error
50	2.95	5.9	1.8	3.6
100	2.95	2.9	2.1	2.1
150	3.97	2.6	3.6	2.4
200	3.96	2.9	3.8	1.9
250	9.65	3.9	8.3	3.3
300	17.33	5.8	12.6	4.2
% Total average		4		2.9

TABLE 7: Comparative Results.

7. CONCLUSION

In this paper, we assess the performance of SIFT, MSER, and SURF, the well known matching algorithms, in solving the correspondence problem and then in estimating the depth within the scene. Furthermore we proposed a framework for estimating the distance between the robot and in front obstacles using the robot stereo camera setup. The results show that the amount of features is not a measure of success by itself but the "quality" of these features. Although the SIFT, detect more matches, but the SURF gives the best result in estimating the depth in all images. We deduce that, the amount of matches detected is not a good indication of the performance of the algorithm. Matches detected by SURF, although fewer, are more robust than those detected by SIFT and MSER. It is concluded that SURF has the best overall performance against SIFT and MSER algorithms. The proposed framework using SURF algorithm performed significantly better than a recent algorithm published, by other researchers, at the same depths. Future work related to this research will be directed to implement SURF algorithm in real time stereo vision navigation and obstacle avoidance for autonomous mobile robot.

8. REFERENCES

1. http://www.surveyor.com/stereo/stereo_info.html
2. Nalpantidis, Lazaros; Gasteratos, A.; Sirakoulis, G.C., "Review of stereo vision algorithms : From software to hardware", International Journal of Optomechatronics, Vol. 2, No. 4, p. 435-462, 2008
3. Dilip K. Prasad, "Survey of the problem of object detection in Real images", International journal of image processing, V (6), issue (6), 2012
4. Manjusha, et al., "A survey of image registration", International journal of image processing, V (5), issue (3), 2011
5. Di Stefano, et al., "A fast area-based stereo matching algorithm", Image and vision computing, 22(12), pp. 983-1005, 2004
6. Harkanwal, et al. , " A robust area based disparity estimation technique for stereo vision applications", proceeding of the 2011 International conference on image processing.
7. Meng Chen, et al., "A method for mobile robot obstacle avoidance based on stereo vision", proceeding 10th IEEE International conference on components, circuits, devices, and systems, 2012
8. Zhao Yong-guo, et al. "The obstacle avoidance and navigation based on stereo vision for mobile robot" proceeding international conference on optoelectronics and image processing, 2010
9. Ibrahim El rube, et al., "Automatic selection of control points for remote sensing image registration based on multi-scale SIFT", proceeding of international conference on signal, image processing, and applications, 2011
10. Ming Bai, et al., "Stereo vision based obstacle detection approach for mobile robot navigation", proceeding international conference on intelligent control and information processing, 2010
11. Patrik Kamencay, et al., "Improved depth map estimation from stereo images based on hybrid method", Radio engineering, vol. 21, NO. 1, pp. 70-78, 2012

12. Sukjune Yoon, et al., "Fast correlation-based stereo matching with the reduction of systematic errors", Pattern Recognition Letters 26, 2221-2231, Elsevier, 2005
13. Kanade, T., and M. Okutomi, " A stereo matching algorithm with and adaptive window: Theory and experiment" IEEE Transactions on Pattern Analysis and Machine Intelligence, (TPAMI) 16: 920-932, 1994
14. J. Matas, O. Chum, M. Urban, and T. Pajdla. "Robust wide baseline stereo from maximally stable extremal regions", Proc. of British Machine Vision Conference, pp 384-396, 2002.
15. Lowe, D.G., "Distinctive image feature from scale-invariant keypoints", International Journal of Computer Vision, 60(2): 91-110, 2004
16. Hess, R. " An open source SIFT library". Proceedings of the International Conference in Multimedia, 2010, Firenze, Italy. pp. 1493-1496, 2010
17. Herbert Bay, H., A. Ess, T. Tuytelaars, and L. Van Gool. "Speeded-up robust features (SURF)", Computer Vision ECCV 2006, Vol. 3951. Lecture Notes in Computer Science. p. 404-417, 2006
18. Evans, C." Notes on the Open SURF library". UK Publication Papers. Issue 1, Citeseer. p. 25., 2008
19. http://www.vision.caltech.edu/bouguetj/calib_doc/index.html#links
20. Luo Juan, Oubong, "A comparision of SIFT, PCA-SIFT, and SURF", International journal of image processing, V (3), issue (4), 2009
21. Young Soo, et al., " Distance estimation using inertial sensor and vision", Inernational Journal of control, Automation, and Systems, V(11), (1), 2013

Performance Evaluation of Object Tracking Technique Based on Position Vectors

V Purandhar Reddy

Associate Professor,
Dept of ECE, S V college of Engineering,
Tirupati, 517507, India.

vpreddy4@gmail.com

K Thirumala Reddy

Assistant Professor,
Dept of ECE, S V college of Engineering,
Tirupati, 517507, India.

kanipakamthirumala@gmail.com

Y.B.Dawood

Dept of ECE, S V college of Engineering,
Tirupati, 517507, India.

dawoodyb@gmail.com

K.M.YaagnaTheja

Dept of ECE, S V college of Engineering,
Tirupati, 517507, India.

yaagnatheja@gmail.com

C.Munesh

Dept of ECE, S V college of Engineering,
Tirupati, 517507, India.

muneshneverstepbacks@gmail.com

M.V.L.Sraavani

Dept of ECE, S V college of Engineering,
Tirupati, 517507, India.

sraavani.mvl@gmail.com

Abstract

In this paper, a novel algorithm for moving object tracking based on position vectors has proposed. The position vector of an object in first frame of a video has been extracted based on selection of region of interest. Based on position vector in first frame object direction has shown in nine different directions. We extract nine position vectors for nine different directions. With these position vectors next frame is cropped into nine blocks. We exploit block matching of the first frame with nine blocks of the next frame in a simple feature space by Discrete wavelet transform and dual tree complex wavelet transform. The matched block is considered as tracked object and its position vector is a reference location for the next successive frame. We describe performance evaluation and algorithm in detail to perform simulation experiments of object tracking using different feature vectors which verifies the tracking algorithm efficiency.

Keywords: Object Tracking, DWT, DTCWT, Feature Vector, Block Matching.

1. INTRODUCTION

The moving object tracking in video pictures has attracted a great deal of interest in computer vision. For object recognition, navigation systems and surveillance systems, object tracking is an indispensable first-step.

The conventional approach to object tracking is based on the difference between the current image and the background image. However, algorithms based on the difference image cannot

simultaneously detect still objects. Furthermore, they cannot be applied to the case of a moving camera. Algorithms including the camera motion information have been proposed previously, but, they still contain problems in separating the information from the background.

In this paper, we propose object tracking in video pictures based on position vectors. Our algorithm is based on position vector calculation and block matching by feature vectors. The proposed method for tracking uses block matching between successive frames. As a consequence, the algorithm can simultaneously track multiple moving and still objects in video pictures.

This paper is organized as follows. The proposed method consisting of stages position vector calculation, feature extraction by different transforms, block matching and minimum distance measure which are described in detail.

2. PROPOSED MOVING OBJECT TRACKING TECHNIQUE

2.1 Extraction of Position Vectors

2.1.1 First Frame position vector extraction

In general, image segmentation and object extraction methods are used to calculate position vectors. In the proposed concept, first select the portion of an object which is to be tracked. The portion of an image is cropped in the first frame which is referred as block.

Based on co-ordinate parameters of an block, we extract the position of the pixel $P_{x_{max}}$ ($P_{x_{min}}$) which has the maximum (minimum) x-component.

$$\begin{aligned} P_{x_{max}} &= (X_{max,x}, X_{max,y}), \\ P_{x_{min}} &= (X_{min,x}, X_{min,y}), \end{aligned}$$

Where $X_{max,x}$, $X_{max,y}$, $X_{min,x}$, and $X_{min,y}$ are x and y coordinates of the

Rightmost and leftmost boundary of the block, respectively. In addition, we also extract

$$\begin{aligned} P_{y_{max}} &= (Y_{max,x}, Y_{max,y}), \\ P_{y_{min}} &= (Y_{min,x}, Y_{min,y}). \end{aligned}$$

Then we calculate the width w and the height h of the block as follows

$$\begin{aligned} w_i(t) &= X_{max,x} - X_{min,x}, \\ h_i(t) &= Y_{max,y} - Y_{min,y}. \end{aligned}$$

We define the positions of each block in the frame as follows

$$\begin{aligned} P &= (X_1, Y_1) \\ X_1(t) &= (X_{max,x} + X_{min,x})/2 \\ Y_1(t) &= (Y_{max,y} + Y_{min,y})/2 \end{aligned}$$

2.1.2 Position Vectors in nine different directions

Frame to Frame movement distance of an object is negligible. So, we consider movement shift by "m" units in nine different directions as shown in the figure (1).

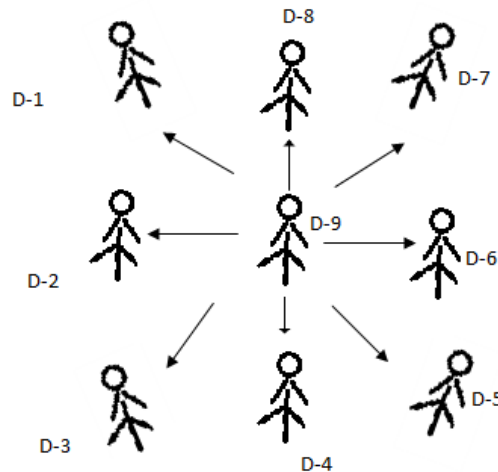


FIGURE 1: Movement of Object in Nine Different Directions.

For Direction 1, position vector shift $P_1 = (X_1 - m, Y_1)$

For Direction 2, position vector shift $P_2 = (X_1 + m, Y_1)$

For Direction 3, position vector shift $P_3 = (X_1, Y_1 + m)$

For Direction 4, position vector shift $P_4 = (X_1, Y_1 - m)$

For Direction 5, position vector shift $P_5 = (X_1 - m, Y_1 - m)$

For Direction 6, position vector shift $P_6 = (X_1 + m, Y_1 + m)$

For Direction 7, position vector shift $P_7 = (X_1 - m, Y_1 + m)$

For Direction 8, position vector shift $P_8 = (X_1 + m, Y_1 - m)$

For Direction 9, position vector shift $P_9 = (X_1, Y_1)$

Based on the position vectors $P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8$ and P_9 crop the second frame into nine blocks.

2.2 Feature Vectors Using DWT & DTCWT

For testing the developed algorithm, Feature vectors are calculated by two methods one is by using Discrete Wavelet Transform (DWT) and the other is Dual Tree Complex Wavelet Transform (DTCWT).

2.2.1 2D-Discrete Wavelet Transform

A weakness shared by most of the texture analysis schemes is that the image is analyzed at one single-scale. A limitation that can be lifted by employing multi-scale representation of the textures such as the one offered by the wavelet transform. Wavelets have been shown to be useful for texture analysis in literature, possibly due to their finite duration, which provides both frequency and spatial locality. The hierarchical wavelet transform uses a family of wavelet functions is associated scaling functions to decompose the original signal/image into different sub bands. The decomposition process is recursively applied to the sub bands to generate the next level of the hierarchy. At every iteration of the DWT, the lines of the input image (obtained at the end of the previous iteration) are low-pass filtered and high pass filtered. Then the lines of the two images obtained at the output of the two filters are decimated with a factor of 2.

Next, the columns of the two images obtained are low and high pass filtered. The columns of those four images are also decimated with a factor of 2. Four new sub-images representing the result of the current iteration are generated. The first one, obtained after two low-pass filtering, is named approximation sub-image (or LL image). The others three are named as detail sub-images: LH, HL and HH. The LL image represents the input for the next iteration.

In this paper we have used level 2 Haar transform. Only the second level LL image is used for the analysis as that contains most of the important information for feature vector calculation. For first frame the feature vector is f_1 based on cropped image using position vector P_1 . For next frame feature vector set is $V=\{V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9\}$ based on position vectors $P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8$ & P_9 .

2.2.2 Dual Tree Complex Wavelet Transform

Dual Tree Complex Wavelet Transform is a recent enhancement technique to the Discrete Wavelet Transform with some additional properties and changes. It is an effective method for implementing an analytical wavelet transform.

DTCWT gives the complex transform of a signal using two separate DWT decompositions i.e., tree 'a' and tree 'b'. DTCWT produces complex coefficients by using a dual tree of wavelet filters and gives real and imaginary parts.

The DTCWT has following properties:

- Approximate shift invariance;
- Good directional selectivity in 2-dimensions (2-D) with Gabor-like filters also true for higher dimensionality: m-D);
- Perfect reconstruction (PR) using short linear-phase filters;
- Limited redundancy: independent of the number of scales: 2:1 for 1-D 2^m :1 for m-D)
- Efficient order-N computation - only.
- DTCWT differentiates positive and negative frequencies and generates

Six sub bands oriented in $\pm 15^\circ$, $\pm 45^\circ$, and $\pm 75^\circ$. The different levels of DTCWT such as levels 5, 6, and 7 are applied on Cropped Blocks.

For performance analysis of algorithm we have used DTCWT Vectors for feature vector calculation. For first frame the feature vector is f_1 based on cropped image using position vector P_1 . For next frame feature vector set is $V=\{V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9\}$ based on position vectors $P_1, P_2, P_3, P_4, P_5, P_6, P_7, P_8$ & P_9 .

2.3 Block Matching and Distance Measure

Using the Feature Vector in the first frame and feature vectors of next frame we perform the minimum distance by Euclidean distance between feature vectors. Finally the cropped image is identified with minimum distance feature vector of next frame. Repeating this matching procedure for all the frames with first frame, we can identify all cropped area one by one and can keep track of cropped area between frames.

2.3.1 Object Tracking Algorithm

1. Input video
2. Crop the first frame for ROI(Region Of Interest)
3. Calculate position vector for cropped portion. $P=(X_1, Y_1)$
4. Based on position vector calculate nine position vectors with X and Y co-ordinate shift by 'm' units. $P_1=(X_1-m, Y_1)$, $P_2=(X_1+m, Y_1)$, $P_3=(X_1, Y_1+m)$, $P_4=(X_1, Y_1-m)$, $P_5=(X_1-m, Y_1-m)$, $P_6=(X_1+m, Y_1+m)$, $P_7=(X_1-m, Y_1+m)$, $P_8=(X_1+m, Y_1-m)$, $P_9=(X_1, Y_1)$
5. For previous frame i.e first frame perform feature vector extraction method for cropped block to get feature vector f_1 .
6. Similarly perform DWT for all cropped blocks in the next frame $K+1$ to get the feature vectors and for vector set V.
7. Block matching distance measure
 - a.) Calculate the distance measured using Manhattan distance between f_1 and V.
 - b.) Apply Feature Vector matching of K^{th} frame cropped image with minimum distance block of $K+1^{th}$ frame. If not matched, perform for next successive frame.
 - c.) After matching remove the position vector data of N^{th} frame and store the data of position vector of $K+1^{th}$ frame.
 - d.) Increase the value of K by $K+1$.
8. Repeat the steps from 1 to 7.

3. IMPLEMENTATION

The implementation of proposed object tracking technique using DWT & DTCWT Feature Vectors is done in MATLAB 7.0.

3.1 Video Data

The object tracking techniques by DWT & DTCWT are tested on the Video Pictures. To compare the techniques and to check their performance we have used the Precision of object Tracking and tracking speed.

3.2 Precision of Object Tracking

Here we address some of the features of an efficient Object tracking system such as accuracy, stability and speed. To assess the object tracking effectiveness, we have used the precision as Statistical comparison parameters for the DWT and DTCWT. The definitions of Precision is given by the equation

$$\text{Object Tracking Precision} = \frac{\text{Number of frames with object tracking Result}}{\text{Number frames given for object tracking test}} \times 100$$

4. RESULTS AND DISCUSSION

The object tracking methods by DWT and DTCWT were applied to the Video frames. The precision is 95% in case of DTCWT and 92 % in case of DWT for 100 frames. This shows that DTCWT outperforms DWT. As far as precision is concerned DTCWT gives best performance.

Table 1 shows the actual precision values computed for considered query frame at percentage precision using two techniques discussed in the paper.



FIGURE 2: Tracking Results from Successive Frames Using DTCWT Feature Vectors.



FIGURE 3: Tracking Results from Successive Frames Using DWT Feature Vectors.

Method	Precision based on number of frames						Execution time between two frames
	5	10	15	20	25	30	
DWT	100%	100%	93.3%	90%	92%	93.3%	0.4msec
DTCWT	100%	100%	100%	95%	96%	96.66%	0.2msec

TABLE 1: Percentage Precision Values for Object Tracking by DWT and DTCWT Feature Vectors.

5. CONCLUSION

The performance of object tracking system depends on the precision and speed. Percentage Precision is taken as criteria for judging the performance of object tracking technique. Thus DTCWT gives the best performance if precision and speed both are considered. Simulation results for frame sequences with moving objects verify the suitability of the algorithm for reliable moving object tracking. We also have confirmed that the algorithm works very well for more complicated video pictures including rotating objects and occlusion of objects. It is obvious that, the simulation result in the proposed algorithm is quite enough to apply for the real time applications. We would like to implement this algorithm with feature vectors in different vectors for future applications too.

6. REFERENCES

1. Moving object tracking based on position vectors V. Purandhar Reddy, International Journal of Scientific & Engineering Research in "Digital Image Processing", volume 4, Issue 4, April 2013.
2. Object Tracking based on pattern matching V. Purandhar Reddy, International Journal of Advanced Research in "Computer Science and Software Engineering", Volume 2, Issue 2, February 2012.
3. Object Tracking based on Image Segmentation and pattern matching V. Purandhar Reddy, International Journal of Computer Science and Information technology, Issue 2, May 2012.
4. H. Kimura and T. Shibata, "Simple-architecture motion-detection analog V-chip based on quasi-two-dimensional processing," Ext. Abs. of the 2002 Int. Conf. on Solid State Devices and Materials (SSDM2002), pp. 240– 241, 2002.
5. John Eakins and Margaret Graham. Content-based image retrieval. Project Report University of Northumbria at Newcastle.
6. Datta et.al. 2005. Content-based image retrieval approaches and trends of the new age. In Proceedings of the 7th ACM SIGMM International Workshop on Multimedia information Retrieval. Singapore: ACM Press. pp.253-262.
7. Datta et. al. 2008. Image retrieval: idea influences, and trends of the new age. ACM Computing Surveys.vol.40(2), pp.1-60.
8. Stricker et. al. 1995. Similarity of color image. In Proceedings of SPIE on storage retrieval for image and video databases, California.Vol. 2420, pp. 381-392.
9. Gao Li-chun and Xu Ye-qiang. 2011. Image retrieval based on relevance feedback using blocks weighted dominant colors in MPEG-7. Journal of Computer Applications.vol.31(6), pp.15491551.
10. H B Kekre and V A Bharadi. Modified BTC & walsh coefficients based features for content based image retrieval. NCICT, India.

11. S.Vidivelli and S.Sathiya devi. 2011. Wavelet based integrated color image retrieval. IEEE-International Conference on Recent Trends in Information Technology, ICRTIT.

12. Kekre et. al. 2010 IEEE. Content based image retrieval using fusion of gabor magnitude and modified block truncation coding. Third International Conference on Emerging Trends in Engineering and Technology.

Comparative Analysis of Partial Occlusion Using Face Recognition Techniques

N.Nallammal

Research Scholar/Dept of computer Science
Avinashilingam Institute for home science
and Higher education for Women
Coimbatore, 641 043, Tamil nadu,India

msg2nalls@gmail.com

Dr.V.Radha

Associate Professor /Dept of computer Science
Avinashilingam Institute for home science
and Higher education for Women
Coimbatore, 641 043, Tamil nadu,India

radharesearch@yahoo.com

Abstract

This paper presents a comparison of partial occlusion using face recognition techniques that gives in which technique produce better result for total success rate. The partial occlusion of face recognition is especially useful for people where part of their face is scarred and defect thus need to be covered. Hence, either top part/eye region or bottom part of face will be recognized respectively. The partial face information are tested with Principle Component Analysis (PCA), Non-negative matrix factorization (NMF), Local NMF (LNMF) and Spatially Confined NMF (SFNMF). The comparative results show that the recognition rate of 95.17% with $r = 80$ by using SFNMF for bottom face region. On the other hand, eye region achieves 95.12% with $r = 10$ by using LNMF.

Keywords: Partial Face Occlusion, Non-Negative Matrix Factorization (NMF), Local NMF, Spatially Confined NMF.

1. INTRODUCTION

Face recognition deals with verifying or identifying a face from its image. It has received substantial attention and its performance has advanced significantly over the last three decades due to its value both in understanding how the face recognition process works in humans as well as in addressing many applications, including access control and video surveillance. For example, individuals who wear sunglasses, masks and veils are restricted to provide their full face due to occupation, contagious disease, privacy and religion practices [1]. Therefore, an automated partial face recognition system based on users' preference could be one of the potential solutions to solve the conflict arouse in [2, 3, 4] previously. For the past few years, researchers [5, 6, 7, 8] had studied on the possibility of using partial face as an alternative for recognition. [6, 7] had applied Radial Basis Function network in symmetric, that examined equal ratio of left side faces and right side faces. Their works had received encouraging results that even achieved the equivalent results of full face. However, the partition of face into left and right side is not practical for direct access control. Meanwhile, [9] had introduced masks on users, where different part of faces are covered and the partial face is evaluated using Lophoscopic Principle Component Analysis. Their algorithm was blemished because it is more computational expensive than Principle Component Analysis (PCA). In this paper, a partial face recognition system framework is compared in well-controlled environments when the full information of face is absent. We examined two scenarios. They are individuals who wear sunglasses and individuals who wear masks or veils during authentication. This means only bottom part of the face and top part of the face especially eye region will be taken into account during recognition.

On the other hand, the term ‘controlled environment’ refers to normal face recognition environment that capture full or partial 2D front face with limited users’ support without intrude users’ privacy. Linear subspace projection has been used extensively for feature extraction in face images. They include PCA [10], Linear Discriminant Analysis (LDA) [11] and neural network approaches [12]. These methods map the high dimensionality images into a lower-dimensional manifold and treat the images as a whole. [15] Proposed Non-Negative Matrix Factorization (NMF) for learning of face features. We made used of NMF and its variants ie. Local NMF [16] and Spatially Confined NMF (SFNMF) [17] to reduce the dimensionality of the raw image and at the same time to preserve as many salient features as possible. In addition, we compare our results with PCA as our baseline. The outline of the paper is organized as follow: Section 2 presents the overview of feature extraction literature. Section 3 is denotes the experimental results and conclusion is discussed in Section 4.

2. FEATURE EXTRACTION LITERATURE

2.1 Principal Component Analysis (PCA)

Turk and Pentland [10] used Principal Component Analysis or known as Eigen Face to represent, detect and recognize faces. Images of faces, being similar in overall configuration, will not be randomly distributed in this huge image space and thus can be described by a relatively low dimensional subspace. The main idea of PCA is to find the vectors that best account for the distribution of face images within the entire image space. These vectors define the subspace of face images, which we call “face space”. Each vector is of length N^2 , describes an $N \times N$ image, and is a linear combination of the original face images. Because these vectors are the eigenvectors of the covariance matrix corresponding to the original face images, and because they are face-like in appearance, they are named as “Eigen Faces”.

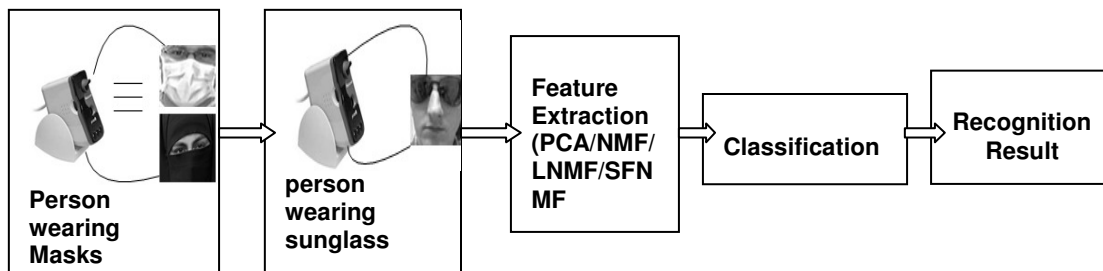


FIGURE 1: Comparison Framework.



FIGURE 2: Examples of AR database.

2.2 Non-Negative Matrix Factorization (NMF)

NMF finds an approximate factorization, where X is the raw face data into non-negative factors W and H . The nonnegative constraints make the representation purely additive (allowing no subtractions), in contrast to many other linear representations such as PCA. This ensures that the components are combined to form a whole in the non subtractive way. Given an initial database expressed by a $n \times m$ matrix X , where each column is an n -dimensional non-negative vector of the original database (m vectors), it is possible to find two new matrices (W and H) in order to approximate the original matrix:

$$X \approx \tilde{X} \equiv WH, \text{ where } W \in \mathcal{R}^{n \times r}, H \in \mathcal{R}^{r \times m} \quad (1)$$

We can rewrite the factorization in terms of the columns of X and H as:

$$x_j \approx \tilde{x}_j = Wh_j, \text{ where } x_j \in \mathcal{R}^m, h_j \in \mathcal{R}^r \text{ for } j = 1, \dots, n \quad (2)$$

The dimensions of the factorized matrices W and H are $n \times r$ and $r \times m$, respectively. Assuming consistent precision, a reduction of storage is obtained whenever r , the number of basis vectors, satisfies $(n + m)r < nm$. Each column of matrix W contains basis vectors while each column of H contains the weights needed to approximate the corresponding column in X using the basis from W . In order to estimate the factorization matrices, an objective function has to be defined. We have used the column of X and its approximation of $X=WH$ subject to this objective function

$$\Theta_{NMF}(W, H) = \sum_{j=1}^n \|x_j - Wh_j\|^2 = \|X - WH\|^2 \quad (3)$$

This objective function can be related to the likelihood of generating the images in X from the basis W and encoding H . An iterative approach to reach a local minimum of this objective function is given by the following rules [18]:

$$W_{ia} \leftarrow W_{ia} \sum_{\mu} \frac{X_{i\mu}}{(WH)_{i\mu}} H_{a\mu} \quad (4)$$

$$W_{ia} \leftarrow \frac{W_{ia}}{\sum_j W_{ja}} \quad (5)$$

$$H_{a\mu} \leftarrow H_{a\mu} \sum_i W_{ia} \frac{X_{i\mu}}{(WH)_{i\mu}} \quad (6)$$

Initialization is performed using positive random initial conditions for matrices W and H . Convergence of the process is also ensured.

C. Local NMF (LNMF)

LNMF [16] aims to improve the locality of the learned features by imposing additional constraints. It incorporates the following three additional constraints into the original NMF formulation.

- (i) LNMF attempts to minimize the number of basis components required to represent X . This implies that a basis component should not be further decomposed into more components.
- (ii) LNMF attempts to maximize the total “activity” on each component. The idea is to retain the basis with the most important information.
- (iii) LNMF attempts to produce different basis as orthogonal as possible, in order to minimize the redundancy between different basis.

LNMF incorporates the above constraints into the original NMF formulation and defines the following constrained divergence as the objective function:

$$\Theta_{LNMF}(W, H) = \sum_i \sum_j X_{ij} \log \frac{X_{ij}}{[WH]_{ij}} - X_{ij} + [WH]_{ij} + \alpha C_{ij} - \beta \sum_i D_i \quad (7)$$

where $\alpha, \beta, > 0$ are constants and $C = W^1W$ and $D = HH^T$. The structure of the LNMF update for W is nearly identical to that in Equation 4, 5; differing only in the coefficient matrix H . The update for H now uses an element-by-element square root to satisfy the three additional constraints:

$$H_{a\mu} \leftarrow \sqrt{H_{a\mu} \sum_i W_{ia} \frac{X_{i\mu}}{(WH)_{i\mu}}} \quad (8)$$

D. Spatially Confined NMF (SFNMF)

SFNMF method is implemented through a series of simple image processing operations to its corresponding NMF basis image. Firstly, a number of *r* original NMF basis are selected. Each basis is processed off-line to detect the spatially confined regions. The maximum values of the basis image are identified by adjust the threshold of a histogram of pixel values and followed by the morphological dilation operation to find a blob region. As a result, SFNMF basis images where only pixels in the detected regions have grey values copied from the corresponding pixels in the original NMF image are created. The remaining pixels are set to zero. SFNMF basis image only represents spatially confined regions. This is intuitive with the idea of recognition by components where spatially confined regions correspond to the important facial features regions such as eyes, eyebrows, nose and lips.

E. Face Recognition In Subspace

As in most algorithms that employ subspace projection, NMF, LNMF, and SFNMF basis are learned from a set of training images. Let α denote the projection vector, the columns of *W* are NMF, LNMF or SFNMF basis images. During recognition, given an input face image, *X*_{test}, it is projected to $\alpha = W^T X_{test}$ and classified by comparison with the vectors 's' α_T that were computed from a set of training images by using the L2 norm distance metric.

3. EXPERIMENTAL RESULTS

For experiment setup, a prototype was developed in MATLAB 7.0, and installed on a 1.60GHz Intel machine with 1Gb of RAM. The experiments are conducted by using Faces-94 Essex University Face Database [19], which consists of 153 subjects with 20 images per person. Face images are of size 180x200 in portrait format and after normalization, it becomes 30x61 for eye region images and 61x73 for bottom face region images. The first 53 subjects with 10 images are used for bases training with a total of 530 images. Another 100 subjects with 20 images are used for testing in the probe set with a total of 2000 images. In our experiments, False Reject Rate (FRR) and False Accept Rate (FAR) tests are performed. A unique measure, Total Success Rate (TSR) is obtained as $TSR = 1 - \frac{FA + FR}{\text{Total number accesses}} \times 100\%$

$$TSR = \left(1 - \frac{FA + FR}{\text{Total number accesses}} \right) \times 100\% . (9)$$

where FA = number of accepted imposter claims and FR =number of rejected genuine claims. For the FAR test, the first image of each subject in the testing set is matched against the first impression of all other faces and the same matching process was repeated for subsequent images, leading to 99,000 (4950 x 20) imposter attempts. For the FRR test, each image of each subject is matched against all other images of the same subject, leading to 19000 (190 attempts of each subject x 100) genuine attempts. An experiment is conducted by using a set of *r*, 2, 4, 6, 8, 10, 20, 40, 60, 80 and 100 to evaluate on the whole face, eye region and bottom face region. PCA with principal component of 100 is used as a baseline for comparison. Table 1 shows the results of face recognition for the user’s whole face, compared with him/her wearing sunglasses, and subsequently a veil/mask. Thus, we are comparing whole face with eye region and bottom face region by adopting PCA, NMF and its variants ie. LNMF ,and SFNMF.PCA acts as a baseline in our study. By using this method, whole face recognition is able to achieve a high TSR of 96.17%. Eye region achieves TSR of 94.09% while bottom face region only 93.47%.Our previous study shows that NMF is slightly inferior than PCA [1, 17, 20, 21]. However, the variants of NMF are robust to performance and the processing time is greatly reduced. LNMF for whole face and eye region are able to achieve the highest TSR of 97.01% with *r* = 60 and 95.12% with *r* = 10 respectively. The bases learned are localized by imposing three additional constraints upon

the original NMF basis [16]. On the other hand, bottom face region achieved the optimum TSR of 95.17% by using SFNMF with $r = 80$. The bases are processed through a series of image processing methods to abolish all noises in an image. Therefore, the basis learnt are said to be more spatially salient and local. Hence, SFNMF demonstrates improvements over NMF and LNMF [21].

Portion		Clean Face	Eye Region	Bottom Face Region
PCA	PC	100	100	100
	FAR	3.82	5.87	6.54
	FRR	3.87	6.13	6.49
	TSR	96.17	94.09	93.47
NMF	r	40	8	20
	FAR	5.57	6.73	7.84
	FRR	5.56	6.98	7.87
	TSR	94.43	93.23	92.16
LNMF	r	60	10	20
	FAR	2.99	4.81	5.50
	FRR	2.98	5.24	5.69
	TSR	97.01	95.12	94.47
SFNMF	r	40	20	80
	FAR	3.30	5.49	4.74
	FRR	3.29	5.69	5.33
	TSR	96.7	94.4	95.17

TABLE 1: Comparison of Partial Occlusion Face Recognition Results Using PCA, NMF, LNMF and SFNMF.

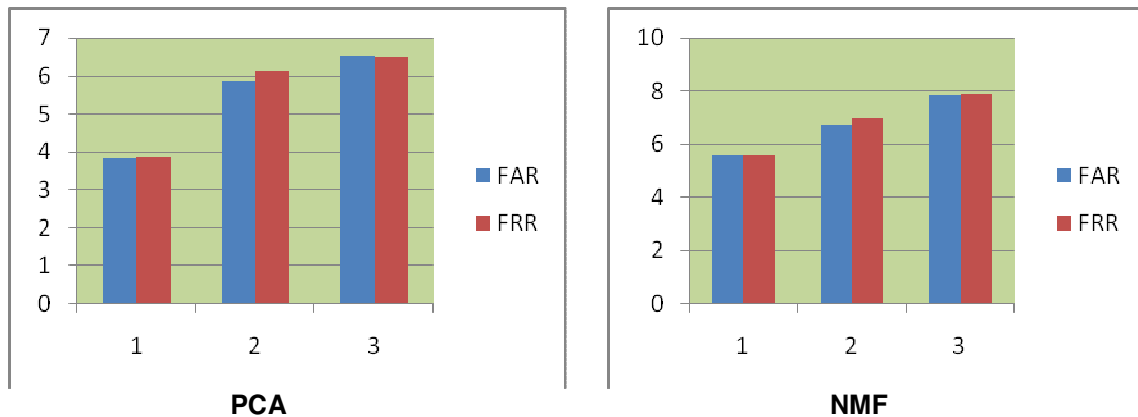


FIGURE 3: Graph Shows Comparison Results of PCA, NMF.

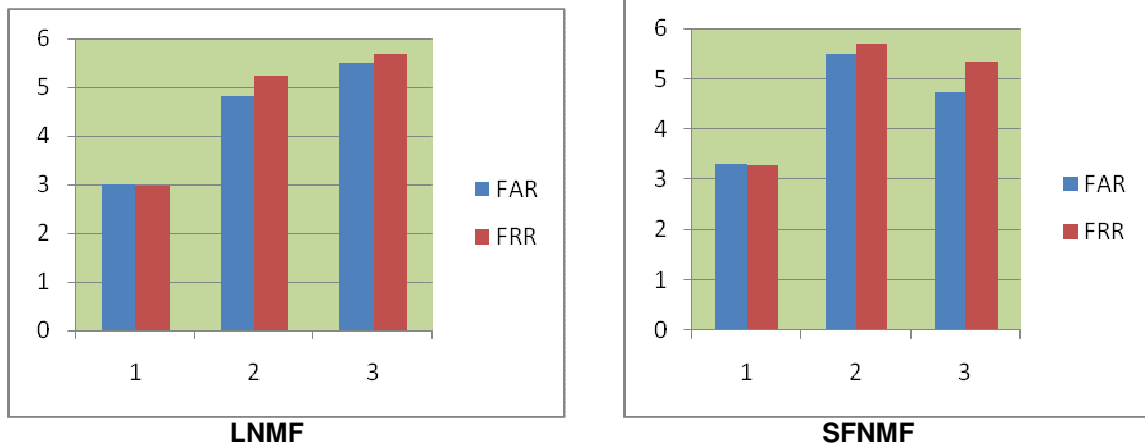


FIGURE 4: Graph Shows Comparison Results of LNMF,SF NMF.

Bottom Region	Total Success Rate
PCA	93.47
NMF	92.16
LNMF	94.47
SFNMF	95.17

TABLE 2: TSR Value of Bottom Region for PCA,NMF,LNMF and SFNMF.

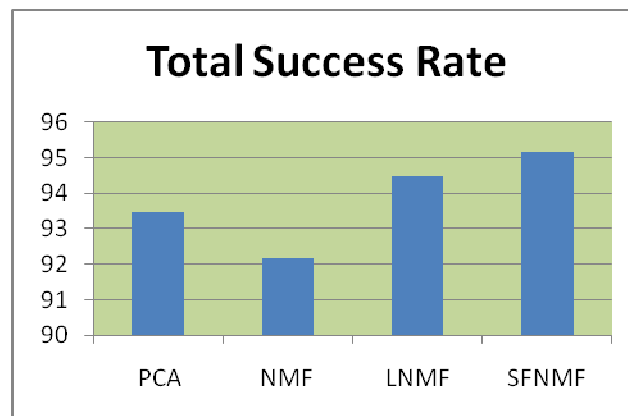


FIGURE 5: Graph Shows Total Success Rate of PCA, NMF,LNMF,SFNMF.

We observe that eye region and bottom face region are able to produce competitive outcomes whereby they only difference by 0.05%. However, the number of r in constructing eye region is only 10 while bottom face region acquires 80. This is due to the image resolution of eye region having smaller size with regard to bottom face region. On top of that, when the bases are locally salient, they require larger r to gain sufficient information to describe a particular face. The graph lies to the axis, the more powerful the recognition system is that SFNMF outperforms NMF and LNMF.

4. CONCLUSION

This paper shows a comparison results for only partial face information ie. Eye region and bottom face region for recognition. Partial face information is crucial in many situations to compensate the absence of full face images. For instance, airports and access control points where some people having facial defects use veils and others down with sickness covered their face with a mask, then our camera would capture only the eye region for authorization. On top of that, for some people wearing sunglasses, we will capture their bottom face region for recognition. Our findings show that bottom face region images itself are achieving a high recognition rate of 95.17% by using Spatially Confined NMF. The result is close to full face TSR of 96.7% of the same method. Therefore, partial face information possesses the unique features thus able to produce fairly good recognition rate. In the future, we would like to improve our feature extraction and classification methods so to increase and surpass the recognition rate of partial face corresponding to whole face.

5. REFERENCES

1. C.C.Teo, H.F.Neo, Andrew B.J.Teoh, "A Study on Partial Face Recognition of Eye Region". International Conference on Machine Vision, pp. 46-49, 28-29 December 2007, Islamabad, Pakistan.
2. Hamdan, Amani, "The Issue of Hijab in France: Reflections and Analysis," Muslim World Journal of Human Rights: vol. 4 : issue. 2,article 4, 2007.
3. Retrieved from <http://www.thesun.co.uk/sol/homepage/news/article76421.ece>, 2006.
4. Retrieved from <http://www.vfs-uk-my.com/biometrics.aspx>, 2005.
5. M. Savvides, R. Abiantun, J. Heo, S. Park, C. Xie and B.V.K. Vijayakumar, "Partial & Holistic Face Recognition On FRGC-II Data Using Support Vector Machine Kernel Correlation Feature Analysis," Proc. Computer Vision and Pattern Recognition Workshop, 2006.
6. S. Gutta, V. Philomin and M. Trajkovic, "An Investigation Into The Use Of Partial-Faces For Face Recognition," Proc. 5th IEEE International Conference on Automatic Face and Gesture Recognition, 2002.
7. S. Gutta, H. Wechsler, "Partial Faces For Face Recognition: Left Vs Right Half," 10th International Conference on Computer Analysis of Images and Patterns, LNCS 2756, pp. 630 – 637, 2003.
8. K.Hotta, "A Robust Object Tracking Method Under Pose Variation and Partial occlusion". IEICE Trans. Inf. & Syst., vol. E89-D, no. 7,pp. 2132 – 2141..
9. F. Tarres, A. Rama, "A Novel Method For Face Recognition Under Partial Occlusion Or Facial Expression Variations," ELMAR 47th International Symposium, pp. 163 – 166, 2005.
10. M. Turk, and A. Pentland, "Eigenfaces for recognition," Journal of Cognitive Neuroscience, vol. 13, no. 1, pp. 71-86, 1991.
11. P.N. Belhumeur, J.P. Hespanha, & D.J. Kriegman, " Eigenfaces vs fisherfaces: recognition using class specific linear projection," Proc. Of European Conf. on Computer Vision. 1996.
12. MJ Er, S Wu, J Lu, HL Toh, Face recognition with radial basis function(RBF) neural network. IEEE Transactions on Neural Networks, vol.13, no.3, pp. 697–710 2002.

13. I. Biederman, "Recognition-by-components: a theory of human image understanding. *Psychological Review*, vol. 94, no. 2, pp. 115-147,1987.
14. Wachsmuth, E., Oram, M.W., & Perrett, D.I., (1994). Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. *Cereb. Cortex*, 22(4):509 - 522.
15. D.D. Lee, and H.S Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, pp. 788-791, 1999.
16. S.Z Li, X.W. Hou, H.J. Zhang and Q. Cheng, "Learning spatially localized, parts-based representation. *IEEE CVPR*, 2001.
17. H.F. Neo, T.B.J. Andrew and N.C.L. David, "A Novel Spatially Confined Non-Negative Matrix Factorization for Face Recognition," *IAPR Conference on Machine Vision Applications*, Tsukuba Science City, Japan. May pp. 16-18, 2005.
18. D.D. Lee, and H.S Seung, "Algorithms for non-negative matrix factorization," *Proceedings of Neural Information Processing Systems*, vol. 13, pp. 556-562, 2001.
19. Vision Group of Essex University Face Database, <http://cswww.essex.ac.uk/mv/allfaces/index.html>, 2004.
20. Han Foon Neo, Chuan Chin Teo, Andrew Beng Jin Teoh, "A Study on Optimal Face Ratio for Recognition Using Part-based Feature Extractor", *Third International IEEE Conference On Signal-Image Technology & Internet-Based Systems (SITIS 07)*, pp.735-741, December 16-19, ShangHai, China, 2007.
21. Andrew Teoh Beng Jin, Neo Han Foon, David Ngo Chek Ling. 2005, "Sorted Locally Confined Non-Negative Matrix Factorization in Face," *IEEE International Conference on Communications, Circuits and Systems (ICCCAS'05)*, Vol. 2, pp. 820-824, 2005.

Image Fusion Quality Assessment of High Resolution Satellite Imagery based on an Object Level Strategy

Farhad Samadzadegan

*Dept. of Surveying and Geomatics University College of Engineering,
University of Tehran,
Tehran, Iran*

samadz@ut.ac.ir

Farzaneh Dadras Javan

*Dept. of Surveying and Geomatics University College of Engineering,
University of Tehran,
Tehran, Iran*

fdadrasjavan@ut.ac.ir

Abstract

Considering the importance of fusion accuracy on the quality of fused images, it seems necessary to evaluate the quality of fused images before using them in further applications. Current quality evaluation metrics are mainly developed on the basis of applying quality metrics in pixel level and to evaluate the final quality by average computation. In this paper, an object level strategy for quality assessment of fused images is proposed. Based on the proposed strategy, image fusion quality metrics are applied on image objects and quality assessment of fusion are conducted based on inspecting fusion quality in those image objects. Results clearly show the inconsistency of fusion behavior in different image objects and the weakness of traditional pixel level strategies in handling these heterogeneities.

Keywords: Image Fusion, Quality Assessment, Object Level, Pixel Level, High Resolution Satellite Imagery.

1. INTRODUCTION

Topographic earth observation satellites, such as IKONOS, Quick Bird and GeoEye, provide both panchromatic images at a higher spatial resolution and multi-spectral images at a lower spatial resolution enjoying rich spectral information [1],[2],[3],[4]. Several technological limitations make it impossible to have a sensor with both high spatial and spectral characteristics [2]. To surmount these limitations, image fusion as a mean for enhancing the information content of initial images to produce new images rich in information content, has drawn an increasing attention in recent years [1],[3]. Remote sensing communities have also switched to merge multi spectral and panchromatic images to exhibit complementary characteristics of spatial and spectral resolutions [2],[5]. This new product is entitled as pan-sharpen image. Nevertheless, as these new images do not exactly have the behavior of the real objects, acquired by remote sensing sensors, quality assessment of these data is crucial before using them in further process of object extraction or recognition. The widespread use of pan-sharpen images has led to a rising demand of presenting methods for evaluating the quality of these processed images [6],[7],[8],[9],[10].

2. IMAGE FUSION QUALITY METRICS (IFQMs)

Image quality metrics are classified based on the level of spectral information [9],[11]. Traditionally, these metrics are classified to mono-modal and multi-modal techniques [12]. A mono-modal metric applies to a single modality while a multi-modal metric applies to several modalities.

Thomas and Wald applied Difference In Variance (DIV), standard deviation and correlation coefficient as mono modal metrics. They applied the metrics for quality evaluation of well-known images of the mandrill and Lenna and images were acquired by satellite observing systems, SPOT-2 and SPOT-5 [11]. Similarly, Riyahi et al., made use of DIV and correlation coefficient as quality metrics to evaluate fusion performance of QuickBird satellite imagery [13]. Chen and Blum, performed some experimental tests according to evaluate quality of image fusion for night vision [14]. They used Standard deviation, SNR (Signal to Noise Ratio) and entropy index as standard quality metrics to extract features from fused image itself. They also used cross entropy based and information based measures to utilize feature of both fused image and source images. Shi et al. applied variety of objective quality metrics, such as correlation, mean value and standard variation, to evaluate wavelet based image fusion of panchromatic Spot image and multi spectral TM image [15].

Entropy, correlation coefficient and mean square error are some of mono modal metrics that were used by Vijarayaji for quantitative analysis of pan-sharpen images [16]. Sahu and Parsai also applied Entropy, SNR and Cross-Correlation to evaluate and have a critical review on recent fusion techniques [17]. Wang et al., introduced the main idea of Structural Similarity (SSIM) which is one of the mono modal metrics. A simplified version of the metric, entitled as Universal Image Quality (UQI) index was introduced by Wang and Bovik (2002) and applied for quality evaluation of IKONOS fused images by Zhang (2008) [8],[18]. Piella and Heijman, 2003, added weighted averaging to UQI to measure the performance of image fusion [7]. This new metric was entitled as saliency factor and was practiced by Hossny et al, for image fusion quality assessment [19]. Piella and Heijman, also introduced weighted saliency factor for fusion quality assessment [7].

On the other hand, Wald introduces ERGAS as a multi-modal index to characterize the quality of process and, present the normalized average error of each band of processed image [6]. Alparone et al., used ERGAS and SAM for image fusion assessment of IKONOS satellite imagery [9]. Riyahi et al., used ERGAS and its modified version RASE (Relative Average Spectral Error) for inspecting different image fusion methods [13]. Van der meer, studied SCM (Spectral Correlation Measure) and SAM for analysis of hyper spectral imagery [20].

Amongst all mono-modal Image Fusion Quality Metrics, UQI has been more frequently used and brought up to be more efficient, reliable and successful [7],[8],[19],[21]. The same story is factual for SAM in terms of multi modal image quality metrics [8],[9],[20]. Our previous results also proved this claim [22].

3. PROPOSED OBJECT LEVEL IMAGE FUSION QUALITY ASSESSMENT

To overcome limitations of the traditional strategies in evaluation of fusion quality with respect to different image objects, this paper presents an object level strategy based on both spectral and shape characteristics of objects (Fig. 1.). In proposed strategy, after generating pan-sharpen image in Phase 1, image objects are extracted from input and pan-sharpen imagery (Phase 2). These objects are computational units for evaluation of fusion quality metrics in phase 3. In phase 4, object level fusion quality assessment is conducted through the whole objects of data set. In the first step, initial panchromatic and multi spectral images are introduced to fusion engine and results in new pan-sharpen image. After generating fused image, the process of evaluating fusion quality based on new strategy is implemented through next three phases. The basic processing units of object-level image fusion quality assessment are image segments, known as image objects, not single pixels. In order to extract image objects, multi resolution image segmentation is carried out in a way that an overall homogeneous resolution is kept. In proposed strategy, based on bottom-up image segmentation, image objects are extracted. In numerous subsequent steps, smaller image objects are merged into bigger ones to minimize average heterogeneity of image objects. The heterogeneity criterion consists of two parts: a criterion for tone and a criterion for shape.

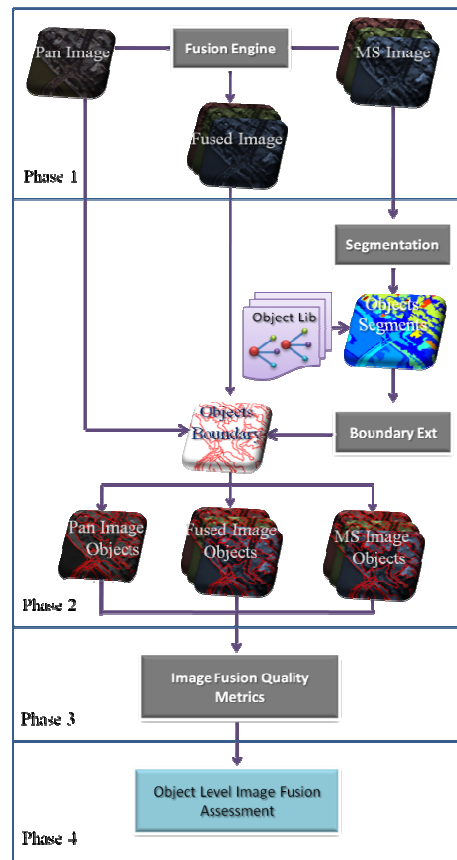


FIGURE 1: Flowchart of Proposed Object Level Fusion Quality Assessment.

When corresponding image objects of all images (panchromatic and multi spectral image and the produced fused image) are determined, image quality metrics are computed for each case. So, quality of corresponding image objects will be inspected. In this study, applied Image quality metrics are SAM and UQI. SAM index is given as:

$$\cos(\alpha) = \frac{\sum_{i=1}^N x_i \cdot y_i}{\sqrt{\sum_{i=1}^N x_i^2 \cdot \sum_{i=1}^N y_i^2}} \quad (3.1)$$

Where N is the number of bands of images or the dimension of the spectral space, $x=(x_1, x_2, \dots, x_N)$ and $y=(y_1, y_2, \dots, y_N)$ are two spectral vectors from the multispectral and fused images respectively [6]. The computed α is the spectral angle for each specific pixel which ranges from 0 to 90 and the minor angle represents the major similarity [6],[9].

On the other hand, Universal Quality Index is computed as:

$$Q = \frac{4 \cdot \sigma_{xy} \cdot \bar{x} \cdot \bar{y}}{(\sigma_x^2 + \sigma_y^2) \cdot (\bar{x}^2 + \bar{y}^2)} \quad (3.2)$$

where \bar{x} and \bar{y} are the local sample means of x and y, σ_x and σ_y are the local sample standard deviations of x and y, and σ_{xy} is the sample cross correlation of x and y after removing their means [18]. Therefore, object level quality assessment will be performed comparing values

of these two metrics. This means, the computational domain of quality evaluation switches from pixel level to object level.

There are two scenarios for object level quality assessment; the type of objects and the effective size of objects in data set. In some applications, the users' purposes about fusion are to make progress and improve the identification potentiality of some specific objects, such as buildings. The quality of these objects should not be less than a specified level of accuracy. In this case, despite the acceptable configuration of general quality of image, fusion process should satisfy a level of quality about specific objects. On the other hand, wide spread objects have more visual effects on pan-sharpen image users. Thus, another object level quality indicator is the evaluation of frequency of image objects pixels against the value of their image quality metric.

4. EXPERIMENTS AND RESULTS

Proposed strategy is implemented and evaluated for quality assessment of high-resolution QuickBird image data over an urban area. The original panchromatic QuickBird has 0.61m pixel while the original multi spectral image has 2.4m pixel spatial resolution (for more information visit digital globe website) [23]. Utilizing PCI software a fused QuickBird image generated with 0.61 meter spatial resolution and three (R,G,B) bands (Fig. 2).

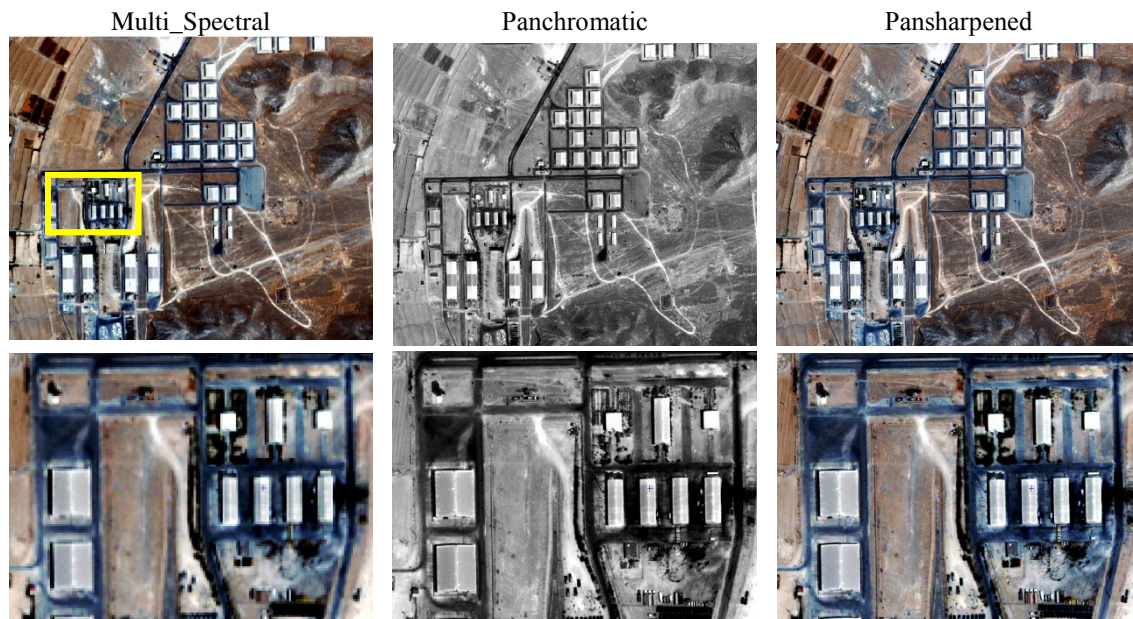


FIGURE 2: QuickBird Dataset.

4.1 Pixel Level Image Fusion Quality Assessment

Pixel level quality assessment of obtained pan-sharpen image is done by computing SAM and UQI statistics for image fusion quality assessment. SAM index is computed for each image pixel of fused image with respect to corresponding multi spectral image pixel, based on equation 3.1. To represent disparity of achieved SAM values, they are represented as pixels intensity values. Achieved image is depicted in Fig. 3.a. By averaging the whole computed SAM indices global measurement of spectral distortion yield and it is presented in Table 1. This final averaged value is what is usually reported as fusion quality in most literatures. Moreover, to have a better perception of fusion behavior, not only the global SAM value, but also the Min, Max and STD values of computed SAM index of all image pixels are presented in Table 1. Moreover, UQI is used to inspect quality of achieved pan-sharpen image as a mono modal metric. This index is computed within a sliding patch with the size of 9 pixels. Final value of UQI is achieved by

averaging computed values of all patches. In order to illustrate UQI behavior, achieved UQI values for each image patch in three layers, R-R, G-G and B-B are averaged and obtained image is depicted in Fig. 3.b.

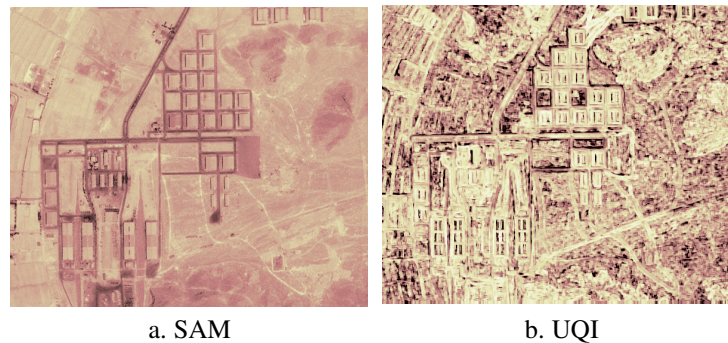


FIGURE 3: Pixel Level Behavior of IFQM Through Data Set.

Moreover, the final value of UQI index, achieved via averaging, and the Min, Max and STD values of achieved UQI in all image patches, are presented in Table 1. Based on the concept of mono modal metrics, they are evaluated for each band of image separately. Consequently, UQI results are presented as the average amount of achieved UQI values for all bands. But, since multi modal metrics treat the image as a 3D data vector and compare the fused image only with the reference multi spectral image, SAM index results are restricted to only one layer. Inspecting results of applying pixel level fusion quality assessment, it is clear that fusion function does not behave uniformly towards the whole image.

TABLE 1: Pixel Level Results of SAM.

Metric	Min/Max	Mean	STD
SAM	0/26	12.56	2.69

It is obvious that the average value for quality metric differs saliently from the min or max values and cannot comprehensively reflect quality of entire image. So, it is an emphasis on non-efficiency of traditional methods of evaluating fusion quality via a single value. Besides, it can be observed that image patches, defined using sliding window for evaluating UQI index, does not match the real image objects and cannot be reliable enough for quality assessment of pan-sharpen image objects. On the other hand, it is obvious that quality values, achieved via each quality metric are completely different. For example in case of SAM it ranges 0-26 while it ranges 0-1 for UQI quality metric. It is realized that there is no individual reference for comparing the outcomes of applying different quality metrics in traditional pixel level fusion quality assessment. All disadvantages of traditional pixel level quality assessment hint to superiority of applying an object level fusion quality assessment for lessening mentioned limitations of traditional pixel level assessment approach.

TABLE 2: Pixel level results of UQI.

Metric	Bands	Min/Max	Mean	STD
UQI	R-R	0/0.89	0.49	0.24
	G-G	0/0.82	0.54	0.18
	B-B	0/0.80	0.49	0.19
	R-P	0/0.82	0.52	0.20
	G-P	0/0.80	0.51	0.20
	B-P	0/0.83	0.45	0.20

4.2 Object Level Image Fusion Quality Assessment

In order to extract image objects, a multi resolution image segmentation method is performed based on the original multi spectral image [24]. For implementation of segmentation, eCognition software system that provides multi resolution object-oriented image analysis is applied (eCognition 4 Professional User guide) [25]. Through the segmentation procedure, the whole image is segmented and image objects are extracted based on adjustable criteria of heterogeneity in color and shape. Achieved segmented image via eCognition software is presented in Fig. 4.a. By implementing image segmentation, different image objects are extracted each of which presents an individual image district. By extracting boundaries of determined image objects and applying them on source panchromatic and pan-sharpen images, corresponding image objects in those imagery are extracted. When image objects extracted, fusion quality is determined for each image object based on SAM and UQI metrics. SAM index evaluated for all pixels of each image object and final value achieved through averaging of all. To show the fusion behavior over image objects, final SAM index for each image object are assigned as pixels intensities and illustrated in Fig. 4.b. On the other hand, in case of UQI, each image segment is considered as image patch, so UQI index achieved for each image object directly applying Equation. 3.2. Average amount of achieved UQI value for all three pan-sharpen image bands with respect to bands of multi spectral image are assigned as pixel intensity values and illustrated in figure. 4.c.

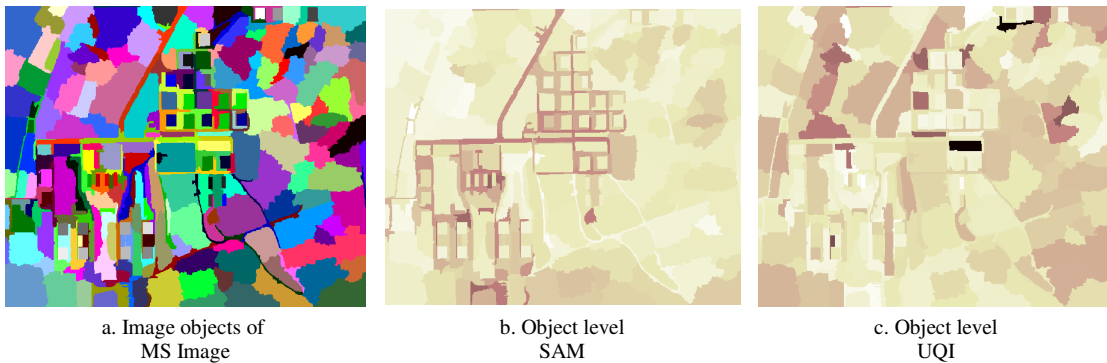


FIGURE 4: Extracted Image Objects and Object Level Behavior of IFQM Through Data Set.

The same as pixel level assessments, the achieved amount of metrics in each individual segment with the Min, Max, Mean and STD values of all segments are determined. Achieved results of both SAM index and UQI are presented in Table 3 and 4.

Table. 2 shows dissimilar statistical behavior of quality index in different image objects for both situation of UQI (mono modal metric) and SAM (multi modal metric).

TABLE 3: Object Level Results of SAM.

Metric	Min/Max	Mean	STD
SAM	3.83/17.07	12.14	2.69

TABLE 4: Object Level Results of UQI.

Metric	Min/Max	Mean	STD	
UQI	R-R	0/0.83	0.59	0.13
	G-G	0.09/0.98	0.92	0.06
	B-B	0/0.82	0.58	0.14
	R-P	0/0.79	0.58	0.12
	G-P	0.50/0.98	0.91	0.07
	B-P	0.50/0.99	0.96	0.04

To assess object level fusion quality, the final results for each metric in all image segments are sorted and visually illustrated to provide better view of fusion behavior (Fig 5). Moreover, to provide a comparative view, all metrics evaluated based on the traditional pixel level strategy and illustrated. Results of applying SAM are presented in Fig 5a. In case of UQI which is a mono modal metric, results are graphically presented in comparison with multi spectral (R-R, G-G, B-B) image (Fig 5.b). The quality metric values achieved traditionally are also plotted.

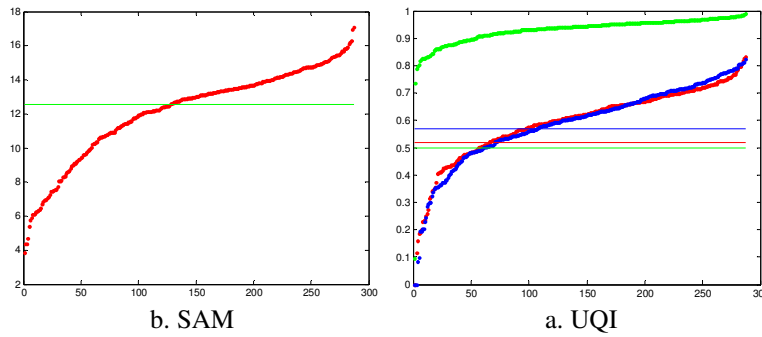


FIGURE 5: Behavior of Object Based IFQMs.

In our experiments, the quality of objects are categorized in three levels, high quality, mean quality objects and low quality objects (Fig 6). Fig 6 shows the frequency of image objects pixels to the value of their image quality metrics of SAM and UQI in the test area.

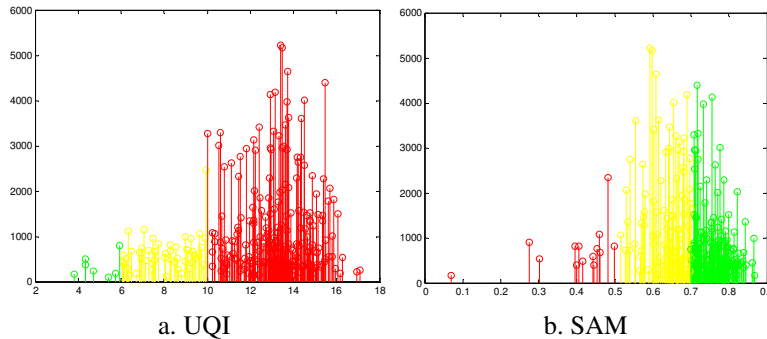


FIGURE 6: Categorization of Fusion Quality in Test Area.

Conducted experiments and obtained results showed that fusion process does not behave uniformly towards the whole image. So, it is not reliable to evaluate fusion quality by a unique quality value. Since for most applications, quality of image objects are of fundamental importance, an object level fusion quality assessment can be helpful in evaluating quality of fusion in different

image objects. Object level quality assessment of fusion lessens the limitations of traditional pixel level strategies. It is also less sensitive to selection of fusion quality metrics.

5. CONCLUSION

There is a wide range of image fusion quality metrics in literature which have been used in different applications and for variety of remote sensing images. In most experiments, these metrics are applied for pixel level fusion quality assessment. This means they evaluate fusion quality in whole image paying no attention to spatial and textural behaviors. This paper proposed an object level fusion quality assessment to model non-uniform behavior of image fusion process. Based on the proposed strategy, image fusion quality assessment is performed for each individual image objects autonomously. Using the high capabilities of this object level image fusion quality assessment strategy, one can solve most of the main problems of traditional pixel level strategies. However, this method still needs some more modifications in the field of definition of image objects which is used in recognition process. Moreover, incorporating of other image quality metrics could efficiently modify the potential of the proposed methodology.

6. REFERENCES

1. D. Fasbender, D. Tuia, P. Bogaert, M. Kanevski, 2008. *Support-Based Implementation of Bayesian Data Fusion for Spatial Enhancement: Applications to ASTER Thermal Images*, IEEE Geoscience and Remote Sensing Letters, Volume: 5 , Issue: 4 Digital Object Identifier: 10.1109/LGRS.2008.2000739 Publication Year: 2008 , Page(s): 598 – 602
2. H. Chu, and W. Zhu, 2008. *Fusion of IKONOS Satellite Imagery Using IHS Transform and Local Variation*. IEEE Geoscience and Remote Sensing Letters, Vol. 5, No. 4
3. F. Bovolo, L. Bruzzone, L. Capobianco, A. Garzelli, S. Marchesi, F. Nencini, 2010, *Analysis of the Effects of Pansharpening in Change Detection on VHR Images*, IEEE Geoscience and Remote Sensing Letters, Volume: 7 , Issue: 1 Digital Object Identifier: 10.1109/LGRS.2009.2029248, Page(s): 53 – 57
4. S. Rahmani, M. Strait, D. Merkurjev, M. Moeller. and T. Wittman, 2010. *An Adaptive IHS Pan-Sharpener*, IEEE Geoscience and Remote Sensing Letters, vol. 7, no. 4, October 2010
5. T. Tu, W. Cheng, C. Chang, P.S. Huang and J. Chang, *Best Tradeoff for High-Resolution Image Fusion to Preserve Spatial Details and Minimize Color Distortion*. IEEE Geoscience and Remote Sensing Letters, Vol. 4, No. 2, April 2007
6. Wald, L., 2000, *Quality of High Resolution Synthesized Images: Is There a Simple Criterion?* Proc. Int. Conf. Fusion Earth Data.
7. G. Piella, H. Heijmans, 2003. *A new quality metric for image fusion*. In: IEEE International Conference on Image Processing, pp. 137–176.
8. Y. Zhang, 2008, *Methods for Image Fusion Quality Assessment-a Review, Comparison and Analysis*, the international archives of photogrammetry, remote sensing and spatial information sciences, Vol XXXVII,(B7). Beijing , China.
9. L. Alparone, S. Baronti, A. Garzelli, and F. Nencini, 2004. *A Global Quality Measurement of Pan-Sharpener Multispectral Imagery*. IEEE Geoscience And Remote Sensing Letters, Vol. 1, No. 4, October 2004

10. S. Li, Z. Lib, J. Gong, 2010, *Multivariate Statistical Analysis of Measures for Assessing the Quality of Image Fusion*, International Journal of Image and Data Fusion Vol. 1, No. 1, March 2010, 47–66
11. C. Thomas and L. Wald, 2006a. *Analysis of Changes in Quality Assessment with Scales*. In Proceedings of FUSION06, 10-13 July 2006, Florence, Italy.
12. C. Thomas and L. Wald, 2006b. *Comparing Distances for Quality Assessment of Fused Products*, In Proceedings of the 26th EARSeL Symposium "New Strategies for European Remote Sensing", 29-31 May 2006, Varsovie, Pologne.
13. R. Riyahi, C. Kleinn and H. Fuchs, 2009. *Comparison of Different Image Fusion Techniques for Individual Tree Crown Identification Using Quickbird Images*, in proceeding of ISPRS Hannover Workshop 2009,
14. Y. Chen, R.S. Blum, 2005. *Experimental Tests of Image Fusion For Night Vision*, in: Proceeding of the 8th International Conference on Information Fusion, 2005
15. W. Shi, Ch. Zhu, Y. Tian, and J. Nichol, 2005, *Wavelet-Based Image Fusion and Quality Assessment*. International Journal of Applied Earth Observation and Geoinformation. v6. 241-251.
16. V. Vijayaraj, 2004. *Quantitative Analysis of Pan-sharpened Images*, Mississippi State University.
17. D. Kumar Sahu and M.P.Parsai, 2012, *Different Image Fusion Techniques –A Critical Review*, International Journal of Modern Engineering Research (IJMER), Vol. 2, Issue. 5, Sep.-Oct. 2012 pp-4298-4301
18. Z. Wang and A. C. Bovik, 2002. *A Universal Image Quality Index*, IEEE Signal Process. Letter, 3 81–4
19. Hossny, M., Nahavandi, Saeid and Creighton, Douglas 2007. A quadtree driven image fusion quality assessment, in *5th IEEE International Conference on Industrial Informatics*, 2007, IEEE Xplore, Piscataway, N.J, pp. 419-424.
20. F. Van Der Meer, 2005. *The Effectiveness of Spectral Similarity Measures for The Analysis of Hyperspectral Imagery*, International Journal Of Applied Earth Observation And Geoinformation, vol. 93, 1-15.
21. Z. Wang, A.C. Bovik, 2009, *Mean Squared Error: Love It or Leave It?* IEEE Signal Processing Magazine, 2009, {26} (1): 98-117
22. F. Samadzadegan and F. DadrasJavan, 2011. *Evaluating the Sensitivity of Image Fusion Quality Metrics*, Journal of Indian Society of Remote Sensing, Published by Springer.
23. QuickBird spacecraft information and specifications, <http://www.digitalglobe.com/index.php/85/QuickBird>
24. U. Benz, P. Hofmann, G. Willhauck, I. Lingenfelder, M. Heynen, 2004. *Multi-Resolution, Object-Oriented Fuzzy Analysis of Remote Sensing Data for GIS-Ready Information*. ISPRS Journal of Photogrammetry & Remote Sensing, 58: 239–258.
25. eCognition 4 Professional User guide <http://www.gis.unbc.ca/help/software/ecognition4/userguide.pdf>.

ISEF Based Identification of RCT/Filling in Dental Caries of Decayed Tooth

A. J. Solanki

Research scholar , EC department,
G. H. Patel college of Engg.& Tech.,
V.V.Nagar, 388120, India

amisolanki007@yahoo.co.in

K. R. Jain

Asst. Prof., EC department,
G. H. Patel college of Engg.& Tech.,
V.V.Nagar, 388120, India

kavindra84@gmail.com

N. P. Desai

Asst. Prof., EC department,
G. H. Patel college of Engg.& Tech.,
V.V.Nagar, 388120, India

niravdesai@gcet.ac.in

Abstract

Dental image processing is one of the emerging fields in case of human identification in forensic sciences. Dental x-rays have been quiet effective for the diagnosis and detection of problems in tooth. This paper presents an add on approach in the same area of medical biometrics to detect and diagnose the dental caries in case of decayed tooth. The enhancement and segmentation of digital dental x-ray image is done by using Infinite Symmetric Exponential filter (Shen Castan Algorithm). The aim of this paper will be to enhance the extracted part of the tooth from digital dental x-ray, finding edges corresponding to caries affected tooth and decide the dental treatment like filling or Root Canal Treatment.

Keywords: Lesion, Enamel, RCT, caries, Dentine, Pulp, ISEF, Dental Radiograph, Dentistry.

1. INTRODUCTION

The process of extracting features, collecting & analysing the useful image information for clinical diagnostics of teeth is the prime need of today's medical science [1]. In this domain of dental image processing, most of the research done is beneficial for forensic science experts for the purpose of human identification. Moving a step ahead in this domain of dentistry the diagnosis of dental diseases from digital dental x-rays is being beneficial and helpful for both doctor as well as patient. Bardia Yousefi et al. in 2012 improved the visibility of digital dental x-ray for teeth, bone and canals using Laplacian transform along with morphological operation. Wavelet transforms and Bayesian classifier is used to classify teeth and canals from resultant image [2]. Ştefan Oprea et al. in 2008 performed dental caries classification based on the edge detection. The dental x-ray image is segmented into individual tooth and then it is converted into binary image of the tooth. The edge detection gives the outline of the dental cavity. The number of caries affected pixels is determined. The caries is classified as *pulpal* if black caries region is adjacent to the white border enclosing the tooth. If there exists two or more number of black regions and the width of the black region is less than 2 mm then it is *Enamel* carry [3].

Prof. G.A. Kulkarni et al. in 2011 proposed two degree differential gray scale method for dental image recognition. The two degree differential method isolated the un-matched part of the two images and gave a satisfied similar rate when the matching location was found. If the matching location was not found, this method enhanced the difference and reduced the similar rate [1]. EyadHaj Said et al. in 2008 performed gray scale stretching transformation for enhancement.

Morphological filtering like top-hat and bottom-hat filters were used for segmentation. 2-D modified wavelet kernels were used to detect boundaries of individual tooth [4].

In this paper we have applied ISEF edge detection to inspect the depth of dental caries in cases of decayed tooth. The paper has been divided in to six parts. Section 2 discusses the problem of dental radiography and possible solution through dental x-ray imaging. Section 3 comprises of the basic concepts of dental caries, the caries affected tooth extraction, its edge detection and further details. In section 4 we propose a novel approach for detection of depth of dental caries. Section 5 concludes the paper. In section 6 acknowledgments are being provided to specialized dental doctors, without their massive support nothing would have been possible.

2. PROBLEM DEFINITION

The raw data obtained directly from x-ray acquisition device may yield a comparatively poor image quality representation. In case of medical images human involvement and perception is of prime importance. It is a difficult task to interpret fine features in various contrast situations.

Nowadays digital dental radiographs, in which enhancement is done automatically, are available but the system are very costly. Our algorithm will give alternate solution to this problem. It includes X-ray imaging & its processing for identifying the exact location & depth of damage in affected tooth. As radiographic imaging study in medical practice provides better clue for diagnosis, but it is not merely the final tool; as investigations must be co-related with clinical findings [Courtesy by Dr. Ronak Panchal].

Dental caries, the common dental diseases, have affected human widely in modern times. Dental caries is an infectious microbiological disease that results in localized dissolution and damage of the calcified tissues of the teeth. Infection of the dental pulp will take place if dental caries are not treated at proper time.

3. BASIC CONCEPTS OF DENTAL CARIES

Classification of dental diseases is decided on the basis of certain criteria, such as based on either the caries lesion is within the enamel, dentine or caries lesion touches the pulp. Dental caries is visible in the x-rays. Image processing techniques will help check the x-rays and detect the depth to which the caries lesion is present and then classify the type of caries present in the dental x-rays. Dental treatment is also dependent on this classification. If caries is developed up to the enamel, it is classified as enamel caries and if caries extended up to the dentine then it is classified as dentinal caries. In above two cases, filling is the best solution. And if caries extended up to the pulp then it is known as pulpal caries, RCT (Root Canal Treatment) is the required treatment.

4. SUGGESTED TECHNIQUE

Sr.No.	Steps
1	Acquire digital dental images.
2	Apply morphological and filtering operations for image enhancement.
3	Extraction of caries affected tooth from image.
4	Edge detection using ISEF (Infinite Symmetric Exponential Filter).
5	Detection and decision based on caries extension inside the tooth as shown in table 4.

TABLE I: Suggested Technique.

4.1 Image Enhancement

In paraipical view, as shown in figure 1, we classify three main classes of “objects”; teeth, gum, and air. An area with “bright” gray scales (except for the pulp tissue) consists tooth area while areas with “mid-range” gray scales consists gum area, and “dark” gray scales indicates air. For better segmentation, it is desirable to convert poor quality dental x-rays in to considerable degree of contrast between the dominant gray scales used in capturing the different classes of objects.



FIGURE 1: Paraipical View of Dental Radiograph
{Courtesy: Dr. Ronak Panchal}.



FIGURE 2: Enhanced Dental X-ray.

Top-hat and bottom-hat filters are applied on the original image to achieve an enhanced and desired image for further processing. Enhanced image is shown in figure 2.

4.2 Caries Affected Tooth Extraction

Caries affected tooth is extracted from enhanced dental X-ray, so that caries affected area can be visible more properly as shown in figure 3.



FIGURE 3: Caries Affected Tooth.

4.3 Edge Detection Using ISEF [16]

Edge detection of caries affected tooth is done by ISEF (Infinite Symmetric Exponential Filter).

Sr.No	Steps
1	Apply ISEF Filter in X direction
2	Apply ISEF Filter in Y direction
3	Apply Binary Laplacian Technique
4	Apply Non Maxima Suppression
5	Find the Gradient
6	Apply Hysteresis Thresholding

TABLE 2: ISEF Algorithm.

Shen Castan Infinite Symmetric Exponential Filter is an optimal edge detector. First the whole image will be filtered by the recursive ISEF filter in X and Y direction respectively which can be implemented by using following equations:

Recursion in x direction:

$$y_1 [i, j] = \frac{(1 - b)}{(1 + b)} I [i, j] + b y_1 [i, j - 1], j = 1 \dots N, i = 1 \dots M \dots \dots (1)$$

$$y_2 [i, j] = b \frac{(1 - b)}{(1 + b)} I [i, j] + b y_1 [i, j + 1], j = N \dots 1, i = 1 \dots M \dots \dots (2)$$

$$r [i, j] = y_1 [i, j] + y_2 [i, j + 1] \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (3)$$

Recursion in y direction:

$$y_1 [i, j] = \frac{(1 - b)}{(1 + b)} I [i, j] + b y_1 [i - 1, j], i = 1 \dots M, j = 1 \dots N \dots \dots (4)$$

$$y_2 [i, j] = b \frac{(1 - b)}{(1 + b)} I [i, j] + b y_1 [i + 1, j], i = M \dots 1, j = 1 \dots N \dots \dots (5)$$

$$y [i, j] = y_1 [i, j] + y_2 [i + 1, j] \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots (6)$$

b=thinning factor (0<b<1)

Subtract the filtered image from the original image to obtain the Laplacian image. In the filtered image, there will be zero crossing in the second derivative at the location of an edge pixel because the first derivative of the image function should have an extreme at the position corresponding to the edge in image. Non maxima suppression is used for thinning purpose for false zero crossing. The gradient is either a maximum or a minimum at the edge pixel. If the second derivative changes sign from positive to negative, it is known as positive zero crossing and if it changes sign from negative to positive, it is known as negative zero crossing. We will permit positive zero crossing to have positive gradient and negative zero crossing to have negative gradient. We considered all other zero crossing as false zero crossing. Thresholding is applied on gradient image. One cutoff is used in simple thresholding but Shen-Castan suggests

for Hysteresis thresholding in which two cut offs are used. Thresholding is applied on the output of an edge detector to decide significant edges. Noise will create spurious response to the single edge that will create a streaking problem. Streaking is defined by breaking up of the edge contour caused by the operator fluctuating above and below the threshold.

Hysteresis thresholding is used to eliminate streaking problem. Individual weak responses usually correspond to noise, but if these points are connected to any of the pixels with strong responses, they are more likely to be actual edge in the image. Such connected pixels are treated as edge pixels if their response is above a low threshold.

We use different thresholds and calculate Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR) values for each experiment. High threshold is selected as 0.8 and low threshold is selected as 0.6. These threshold values gives minimum mean square error as expressed in table 3.

$$MSE = \frac{1}{M * N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} (f(i, j) - g(i, j))^2 \tag{7}$$

Where f(i,j) is the input image and g(i,j) is the edge detected image.

$$PSNR = 10 * [[\log(255 * 255)] / MSE] \tag{8}$$

	Low Threshold				High Threshold			
THRESHOLD	0.2	0.4	0.5	0.6	0.6	0.7	0.8	0.9
MSE	0.3544	0.3618	0.3467	0.3440	0.3622	0.3606	0.3491	0.3558
PSNR	52.6359	52.5463	52.7309	52.7650	52.5409	52.5610	52.7008	52.6185

TABLE 3 Threshold Selection.

The ISEF algorithm is given in table 2. Output is shown in figure 4.



FIGURE 4: ISEF Operated Tooth (LT=60% of max. intensity HT=80% of max. intensity).

5. RESULTS AND DISCUSSION

After converting the dental x ray image to gray scale image, morphological operation using top hat and bottom hat transformation is applied. The resultant image so achieved is shown in figure 5.



FIGURE 5: After Applying Morphological Operation.

Morphological operated image is converted in to binary image. The caries affected tooth is extracted using image cropping operation. Due to presence of noise factor in the binary image median filter is applied on it. The result is shown in figure 6.



FIGURE 6: Removal of Noise Using Median Filtering.

Edge detection of tooth is done using ISEF as shown in figure 4. Now to decide the treatment planning, the following technique is used as presented in Table 4.

Sr.No.	Steps
1.	Decide a horizontal region of interest as shown in Figure - 8.
2.	Bottom line of the region of interest is taken as a reference.
3.	If caries areas are in the enamel region or extended up to the dentine region then filling is to be done.
4.	If caries areas are falling below the dentine region then RCT is required.

TABLE 4: Steps for Deciding Filling / RCT for Horizontal ROI.

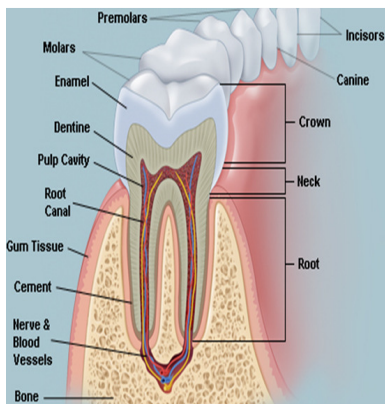


FIGURE 7: Tooth Structure {Courtesy: Dr. Ronak Panchal}.

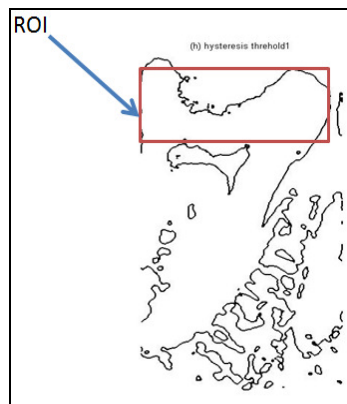


FIGURE 8: Image With Horizontal Region of Interest.

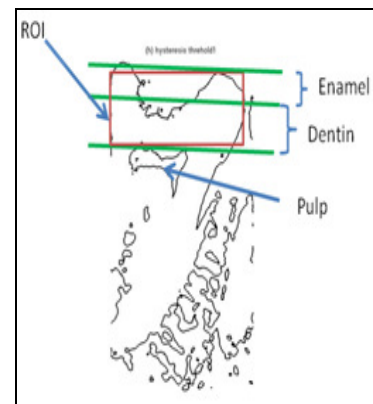


FIGURE 9: Image Showing Different Areas in Horizontal ROI.

Sr.No.	Steps
5.	Decide a vertical region of interest as shown in Figure - 10.
6.	Inner Slanted line of the region of interest is taken as a reference.
7.	If caries areas are in the ROI region then filling is to be done.
8.	If caries areas are falling beyond the ROI region (inner part of the tooth) then RCT is required.

TABLE 5: Steps for Deciding Filling / RCT for Vertical ROI.

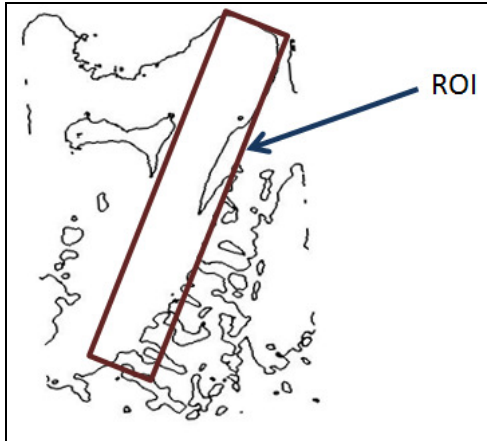


FIGURE 10: Image With Vertical Region of Interest.

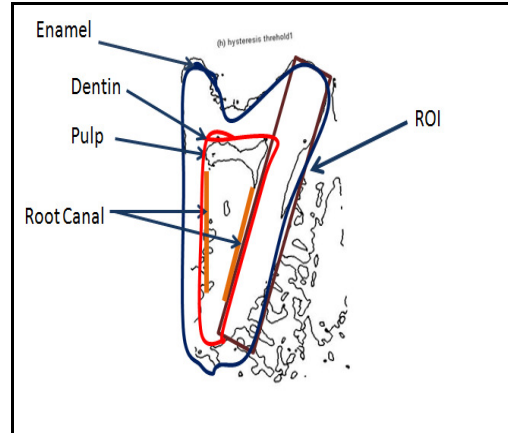












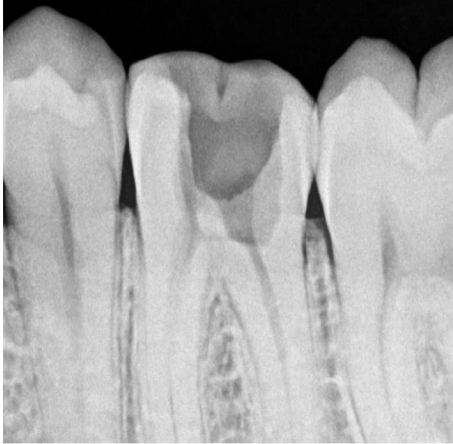





FIGURE 11: Image Showing Different Areas in Vertical ROI.







As discussed in table 4 if caries affected area is on the top side of the tooth in the selected region of interest as shown in figure 8 then filling is to be done. If caries affected area is not only in enamel and dentine region as shown in figure 9 then RCT is to be done.

As discussed in table 5 if decaying is from the sides of the tooth then the selected ROI as shown in figure 10 can be effectively used. Taking figure 7 as reference figure which would give us the idea about how much is the caries affected portion. If the decaying is reached up to the roots as shown in figure 11 then RCT is the only option. The same approach is applied on various dental x-ray images and the result is shown in figure 12.

Sr.No.	Image	Result using reference
1	<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="text-align: center;"> <p>Original image</p>  </div> <div style="text-align: center;"> <p>(b) hysteresis threshold</p>  </div> </div>	RCT-Fig.8

<p>2</p>	<p>Original image</p>  <p>(h) hysteresis threshold</p> 	<p>Filling- Fig. 8</p>
<p>3</p>	<p>Original image</p>  <p>(h) hysteresis threshold</p> 	<p>RCT- Fig.8 and Fig. 10</p>
<p>4</p>	<p>Original image</p>  <p>(h) hysteresis threshold</p> 	<p>RCT- Fig. 10</p>
<p>5</p>	<p>Original image</p>  <p>(h) hysteresis threshold</p> 	<p>RCT- Fig. 8</p>

<p>6</p>	<p>Original image</p>  <p>(b) hysteresis threshold</p> 	<p>RCT- Fig. 8</p>
<p>7</p>	<p>Original image</p>  <p>(b) hysteresis threshold</p> 	<p>Filling- Fig.8</p>
<p>8</p>	<p>Original image</p>  <p>(b) hysteresis threshold</p> 	<p>RCT- Fig. 10</p>

<p>9</p>	<p>Original image</p>  	<p>RCT- Fig. 8 and Fig. 10</p>
<p>10</p>	<p>Original image</p>  	<p>Filling- Fig. 8</p>
<p>11</p>	<p>Original image</p>  	<p>RCT – Fig. 10</p>

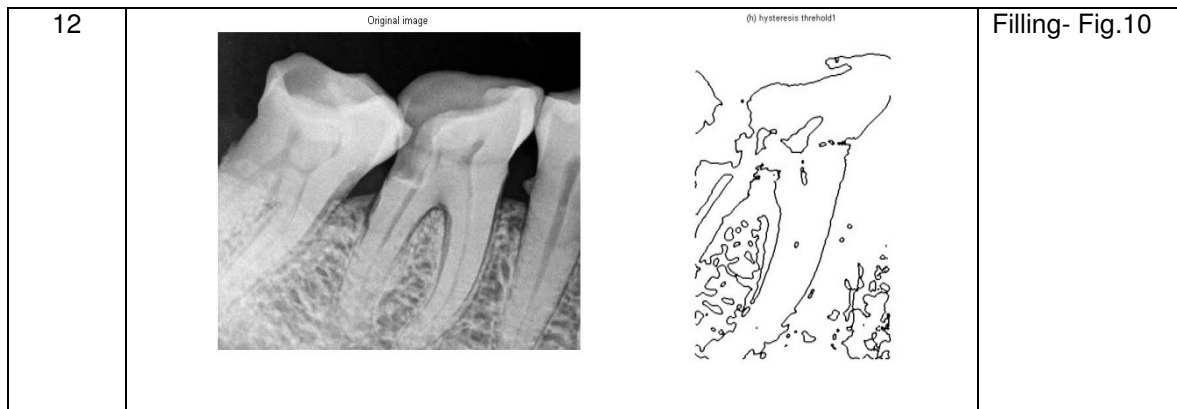


FIGURE 12: Results of Various Dental x-rays of Various Patients for Decision Making to be Operated for Filling or RCT.

6. CONCLUSION

Figure 8, 9, 10, 11 and 12 are the resultant images which were also shown to various dentists and the results were not only accurate but well appreciated and recommended for further research. Figure 12 is the edge detected dental x ray images of patients which gives a clear idea of the caries areas affected either from top or sides. The decision so taken were easy and clear as the dental x-ray edge detected images were noise free as well as clear lines depicted the caries affected areas of enamel, dentine or pulp. In figure 12 second figure (12 (b)) from first row results in filling rest all five results show that RCT has to be done as the carries affected areas have crossed the horizontal or vertical ROI.

7. ACKNOWLEDGEMENT

The authors are highly indebted and thankful to Dr. Ronak Panchal (Shakti Orthodontic Care and Multispeciality Dental Clinic, V.V.Nagar, Anand) for their massive supports and suggestions. We are really grateful to Dr. Ronak Panchal for sharing the personal information of various patients for the purpose of our research. The data base so provided is confidential and presented only after the prior permission of doctor and patients.

8. REFERENCES

1. G.A.Kulkarni, A.S.Bhide, D.G. Patil, S.S.G.B.C.O.E. & T., Bhusawal, "Two Degree Greyscale Differential Method for Teeth Image Recognition", International Journal of computer Application, 2012.
2. B. Yousefi, H. Hakim, N. Motahir, P. Yousefi1, M. M.Hosseini, "Visibility Enhancement of Digital Dental X-Ray for RCT Application Using Bayesian classifier and Two Times Wavelet Image Fusion", Journal of American Science, pp 7-13, 2012.
3. Ş. Oprea, C. Marinescu, I. Liță, M. Jurianu, D. A. Vişan, I. B. Cioc, "Image Processing Techniques used for Dental X-Ray Image Analysis", Electronics Technology, ISSE 2008, pp 125-129.
4. E. H. said, G. Fahmy, D. nassar, H. Amar, "dental X-ray image segmentation" Biometric Technology for Human Identification, Proceedings of the SPIE, Vol. 5404, pp. 409-417, 2004.
5. E. H. Said, D. E. M. Nassar, G. Fahmy, H. H. Ammar. "Teeth segmentation in digitized dental X-ray films using mathematical morphology," IEEE Transactions on information forensic and security, vol. 1, Issue. 2, pp. 178-189, June. 2006.

6. S. Jadhav, R. Shriram, "Dental biometrics used in forensic science", *Journal of Engineering Research and Studies*, Vol.3, pp 26-29, January-March, 2012.
7. S. Dighe, R. Shriram, "Pre-processing, Segmentation and Matching of Dental Radiographs used in Dental Biometrics", *International Journal of Science and Applied Information Technology*, Volume 1, No.2, pp 52-56, May – June 2012.
8. D. B. Prajapati, N.P. Desai, C.K.Modi, "A simple and novel CBIR technique for features extraction using AM dental radiographs", *International Conference on Communication Systems and Network Technologies*, pp 198-202, 2011.
9. R. B. Tiwari, Prof. A. R. Yardi, "Dental x-ray image enhancement based on human visual system and local image statistics", *International Conference on Image Processing, Computer Vision and Pattern Recognition*, 2006, pp 100-108.
10. N. E. Mekky, F. E.-Z. Abou-Chadi, S. Kishk, "Wavelet-Based Image Registration Techniques: A Study of Performance", *International Journal of Computer Science and Network Security*, VOL.11 No.2, pp 188-196, February 2011.
11. P. Ramprasad, H. C. Nagaraj, M. K. Parasuram, "Wavelet based Image Registration Technique for Matching Dental x-rays", *World Academy of Science, Engineering and Technology* 44, 2008, pp 467-470.
12. M. Omanovic, J. Orchard, "Exhaustive Matching of Dental X-rays for Human Forensic Identification", *Journal of the Canadian Society of Forensic Science*, 41(3), pp 1-11, 2008.
13. E. B.Barboza, A. N. Marana, " A Multibiometric Approach in a Semi Automatic Dental Recognition Using DIFT Technique and Dental Shape Features", *SIBGRAPI*, August 22-25, 2012, Ouro Preto, Brazil.
14. S.L. S. Abdullaha, H.A.Hambalia, N. Jamilc, "Segmentation of Natural Images Using an Improved Thresholding-based Technique", *International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012)*, pp 938-944.
15. S. Kiattisin, A. Leelasantitham, K. Chamnongthai, K. Higuchi, " A Match of X-ray Teeth Films Using Image processing Based on Special Features of Teeth" , *SICE Annual Conference 2008*, pp 35-39.
16. C K Modi , K J Pithadiya , J D Chauhan, K R Jain, " Comparative study of Optimal edge detection algorithms for liquid level inspection in Bottles", *International conference on Emerging Trends in Engineering and Technology* , pp 447-452, 2009.
17. M.H.Mahoor, M. A. Mottaleb."Classification and numbering of teeth in dental bitewing images", *Pattern Recognition* 38 (2005), pp 577–586.
18. O. Nomir, M. A. Mottaleb."Hierarchical contour matching for dentalX-ray radiographs", *Pattern Recognition* 41 (2008), pp 130 – 138.
19. O. Nomir, M. A. Mottaleb. "A system for human identification fromX-ray dental radiographs," *Pattern Recognition* 38 (2005), pp 1295 – 1305.
20. O. Gormeza, H. H.Yilmazb," Image Post Processing in Dental Practice", *European Journal of Dentistry* ,October 2009 - Vol.3.

21. S. Shah, A. Abaza, A. Ross, H. Ammar." Automatic Tooth Segmentation Using Active Contour Without Edges", In Biometric Consortium Conference, 2006 Biometrics Symposium: Special Session on Research at the, pp. 1-6. IEEE, 2006.

22. M. Analoui. "Radiographic image enhancement. Part I: spatial domain techniques", *Dento maxillofacial Radiology* (2001) 30, pp 1-9.

Recognition of Facial Expressions using Local Binary Patterns of Important Facial Parts

Ramchand Hablani

*Computer Science and Engg
Sanghvi Institute of Management & Science
Indore 452010 (M.P.) India*

ram.hablani@gmail.com

Narendra Chaudhari

*Computer Science and Egg
Indian Institute of Technology
Indore 452017 (M.P.) India*

nsc183@gmail.com

Sanjay Tanwani

*School of Computer Science & IT
Devi Ahilya University,
Indore 452001 (M.P.) India*

sanjay_tanwani@hotmail.com

Abstract

Facial Expression Recognition is one of the exciting and challenging field; it has important applications in many areas such as data driven animation, human computer interaction and robotics. Extracting effective features from the human face is an important step for successful facial expression recognition. In this paper we have evaluated Local Binary Patterns of some important parts of human face, for person independent as well as person dependent facial expression recognition. Extensive experiments on JAFFE database are conducted. The experiment results show that person dependent method is highly accurate and outperform many existing methods.

Keywords: Facial Expressions, Local Binary Pattern (LBP), Histogram.

1. INTRODUCTION

Facial expression is one of the powerful and natural mean for human beings to communicate their emotions and intentions [1]. Facial expression carries crucial information about the mental, emotional and even physical state of a human being. It is a desirable feature of next generation computers, which can recognize facial expressions and responds accordingly and enables better human machine interactions.

Automated Facial Expression Recognition (AFER) is an interesting and challenging problem. Facial Expression Recognition requires both extraction of facial features and design of a classifier, as shown in figure1.

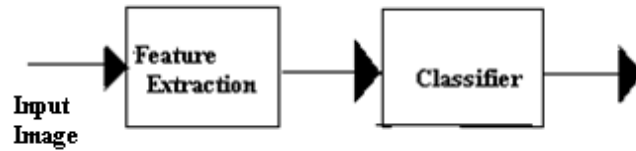


FIGURE 1: Facial Expression Recognition System.

Features (Real values) are extracted from the original face images which minimizes the within class variation of expression and maximizes the between classes variations. If improper features are used, even the best classifier could not recognize proper expressions. There are two main types of approaches to extract facial features: [22-23] geometric feature based methods [2-4] and the appearance based methods [5-9] [16-21]. Geometric feature based methods extract geometric information from the facial images. In appearance based methods, features are either extracted from the entire face or specific regions in facial images. Because of more effectiveness, we are choosing appearance based approach. Gabor wavelet appearance features were demonstrated to be more effective than geometric features [5]. However Gabor Wavelet representation is computationally expensive.

In this paper we make use of facial expression representation based on Local Binary Pattern (LBP) [8, 9, 14] [16-18]. LBP features were proposed originally for texture analysis [6, 7]. Ahonen et al [10, 11] presented LBP based methods for face detection and recognition. The most important properties of LBP features are their tolerance against illumination changes and their computational simplicity.

2. LOCAL BINARY PATTERNS (LBP)

LBP features were originally proposed for texture analysis, which have been recently used in face recognition and facial expression recognition due to its low computation and high discrimination capability. The original LBP operator labels the pixel of an image by thresholding the 3X3 neighborhood of each pixel with the value of the central pixel, and a binary value is assigned to neighborhood pixel on basis of the following function.

$$f(nh) = \begin{cases} 1, & \text{if } v(nh) \geq v(c) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Where $v(nh)$ is gray- scale value of the neighborhood pixel and $v(c)$ is gray- scale value of the centre pixel. These neighborhood bits form a Local Binary Pattern (LBP) corresponding to central pixel.

This can be understood from an example. Suppose the values of a pixel and its eight neighbors are as follows.

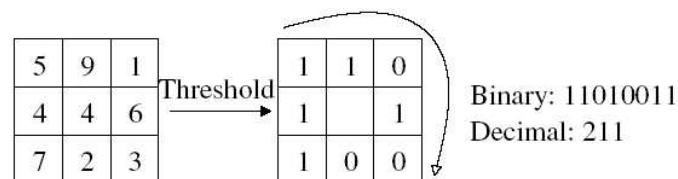


FIGURE 2: Local Binary Pattern.

The derived numbers (called LBP) represent different local patterns like edges, curves, flat regions and spots etc. Using LBP operator the whole image can be transformed to LBP image. An example of LBP image of a facial image is shown in figure 3.

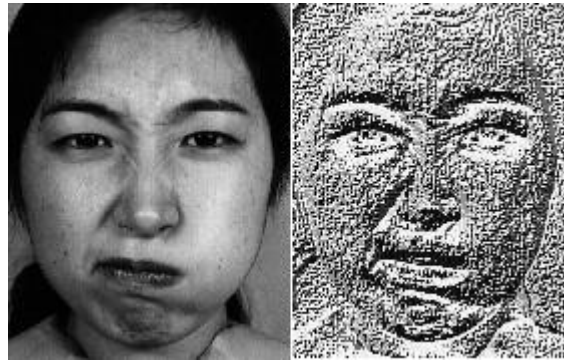


FIGURE 3: An Example of LBP Image for a Facial Image.

The limitation of original binary pattern is its small 3X3 neighborhood, which cannot capture the dominant features. The basic LBP Operator was extended to the neighborhood of different sizes [12]. Using circular neighborhood and bilinear interpolation, the neighborhood of any radius with different number of pixels can be used. See Figure 4 for extended LBP operator.

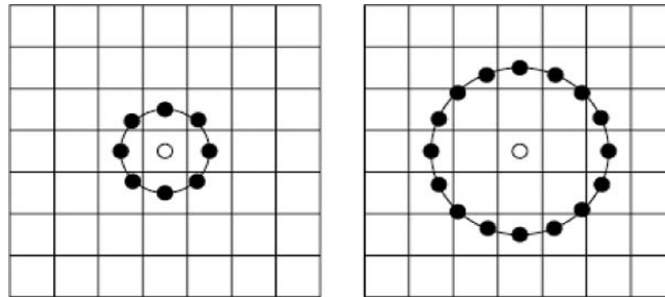


FIGURE 4: Local Binary Patterns of Circular Radius 1 and 2 with 8 and 16 Pixels.

The notation $LBP_{P,R}$ is used to denote the Extended LBP with P pixels and R radius. It has been shown that certain patterns contain more information than others [12]. Therefore it is advantageous to use only those patterns which contain more information. Ojala et al [12] called these patterns as uniform patterns. A Local Binary Pattern is called uniform if it contains at most two bitwise transitions either from 0 to 1 or from 1 to 0. For example, 01111111 is a uniform pattern but 10001101 is not, as it has three bitwise transitions. It is observed that about 90% of the patterns in (8, 1) neighborhood and about 70% of the patterns in (16, 2) neighborhood are uniform patterns in texture images [12]. The LBP operator that accumulates only uniform patterns is denoted by $LBP_{P,R}^{U2}$. The number of patterns for $LBP_{8,1}^{U2}$ is only 59 as compared to number of patterns for standard $LBP_{8,1}$, which is 256.

After applying LBP operator to each pixel of an image, the Histogram of LBP operator values is formed as follows.

$$H_i = \sum_{x,y} I(LBP(x,y) = i) \quad i = 0,1, \dots, n - 1 \quad (2)$$

Where n is the different possible values (Labels) produced by the LBP operator, and

$$I(x) = \begin{cases} 1 & \text{if } x \text{ is True} \\ 0 & \text{if } x \text{ is False} \end{cases} \quad (3)$$

This LBP histogram of an image contains the information about local micro patterns like edges, curves, flat regions and spots etc present in the image. LBP histogram is shown in figure 5.

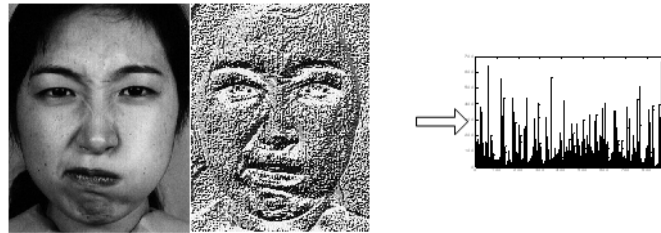


FIGURE 5: LBP Histogram of a Facial Image.

3. FEATURE EXTRACTION FROM FACIAL PARTS

LBP histogram computed from the whole face image tells about the occurrences of the micro-patterns without any indication of their locations. An alternative of this is as follows: first divide the whole face image into sub regions, find out the histogram of each sub region and concatenate all histograms to get a LBP histogram of face. We know that each sub region does not contain the equal amount of information about facial expressions, so we choose some important parts of the face for above purpose. We have chosen eight important sub regions of a face as shown in figure 6, those are two parts of left eye, two parts of right eye, two parts of nose and two parts of mouth. After experimenting with different LBP operators, we have chosen $LBP^{U2}_{8,2}$. The number of bins for each region is 59, so the size of a final feature vector is 472.

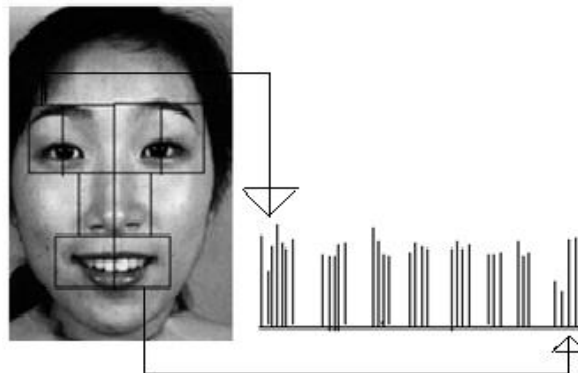


FIGURE 6: Important Facial Parts and Their Histograms of LBP's.

4. FACIAL EXPRESSION RECOGNITION USING LBP

We evaluate the performance on JAFFE (Japanese Female Facial Expression) [13]. JAFFE is a very popular database for facial expression recognition, in which in total 213 facial expression images with 10 Japanese women are involved. Each individual has three or four images with seven kinds of facial expressions, including anger, disgust, fear, happy, sadness, surprise and neutral,. Figure 7 shows seven expression image examples selected from JAFFE database.

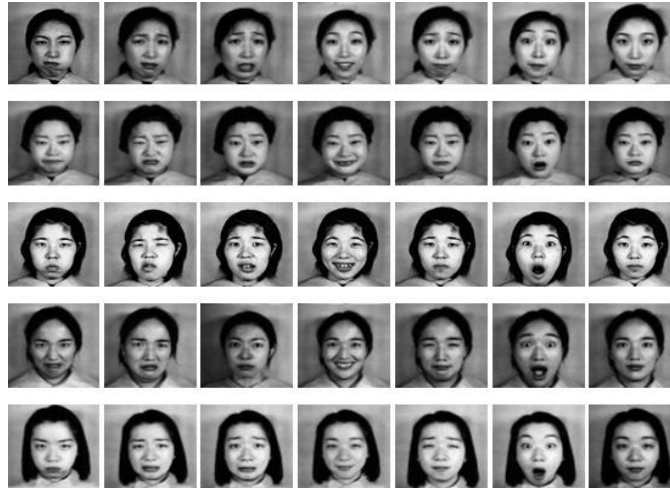


FIGURE 7: Some Sample Images from JAFFE Dataset.

In this section, we perform person-independent as well as person dependent facial expression recognition using LBP features along with template matching as a classifier. Template matching was used in [14] to perform face recognition using LBP-based features: a template is formed for each class of face, and then the nearest neighbor classifier is used to match the test image with the closest template. Here we adopt the template matching for classification of facial expressions. We adopt two types of template matching, one is the person independent template matching and the other is person dependent.

4.1 Person Independent Template Matching

A template is formed for each class of face expression by averaging the LBP histograms of a particular expression. In the training phase, these seven templates are stored. In the testing phase, a test image is compared with all stored templates. We have selected the Chi square test (χ^2) as similarity measure.

$$\chi^2(\mathbf{S}, \mathbf{M}) = \sum_i (S_i - M_i)^2 / (S_i + M_i) \tag{4}$$

Where \mathbf{S} and \mathbf{M} are two LBP histograms of template and test images respectively. Person independent template matching has achieved the generalization performance of 73.61% for 7 category classification. The Confusion matrix for 7-class facial expression recognition is shown in Table1.

Table1: Confusion Matrix for 7-class Person Independent Facial Expression Recognition.

	Anger %	Disgust %	Fear %	Happy %	Neutral %	Sad %	Surprise %
Anger	60	10	0	10	10	10	0
Disgust	0	50	10	20	10	10	0
Fear	0	0	54.5	9.1	9.1	9.1	18.2
Happy	0	0	0	91	9	0	0
Neutral	0	0	0	0	90	10	0
Sad	0	0	10	0	10	80	0
Surprise	0	0	10	0	10	0	80

Note that happy, neutral, sad and surprise expressions can be recognized with high accuracy

(about 80-90%), but anger, fear and disgust are easily confused with other expressions.

4.2 Person Dependent Template Matching

Facial expressions may be expressed differently by different people [15], so low accuracy is achieved in the above method. Therefore, we propose a method that is person dependent. Instead of developing person independent seven templates- one for each expression- we propose to form templates that are person dependent. For each person, seven templates are formed- one for each expression- so total of 70 templates for 10 persons are formed. These 70 templates are stored in training phase. In the testing (Recognition) phase a test image is compared with all stored templates. The one with minimum distance is declared as a recognized expression. The person dependent template matching has achieved a very high generalization performance of 94.44% for 7 category classification. The Confusion matrix for person dependent facial expression recognition is shown in Table2.

Table2: Confusion Matrix for 7-class Person Dependent Facial Expression Recognition.

	Anger	Disgust	Fear	Happy	Neutral	Sad	Surprise
Anger	90	0	0	0	0	10	0
Disgust	0	100	0	0	0	0	0
Fear	0	0	90	0	0	10	0
Joy	0	0	0	100	0	0	0
Sad	0	0	0	0	100	0	0
Surprise	0	0	10	0	0	90	0
Neutral	0	0	0	0	0	10	90

Another benefit of this approach is that along with the recognition of an expression, it also recognizes with expression belongs to which particular person. The time taken to recognize an expression with our approach is, on an average, 11.5 milli seconds.

Person dependent template matching has achieved an average performance of 94.44% for JAFFE database, and has outperformed other methods as listed in Table3.

Table3: Comparison Between Different Methods for 7-class Recognition.

Method(features + classifier)	Recognition Rate (%)
LBP +Template Matching [1]	79.1
Geometric Features +TAN[24]	73.2
LDA+NN[1]	73.4±5.6
LBP+SVM(RBF)[1]	88.9±3.5
Gabor +SVM(RBF)[1]	86.8±3.6
Proposed method(Person Independent)	73.61
Proposed method(Person Dependent)	94.44

5. CONCLUSION AND FUTURE WORK

In this paper, we have extracted features based on Local Binary Pattern. As every part of the face does not contribute equally in face expression recognition, we have chosen some important facial parts like sub parts of eyes, nose and mouth. With the templates of extracted facial features, template matching was used to classify the expression. Experimental results show that the proposed approach is better than approaches that use the whole face image. The proposed method integrates person identity to perform better than conventional expression systems. The

proposed person dependent approach achieves higher recognition rates than those of other approaches. Chi square distance is used as measure of similarity; for classification, we will use neural network and Support Vector Machines (SVM). Manual detection of face and its important parts will be enhanced by automatic detection. Performance evaluation will be extended from JAFEE to other databases.

6. REFERENCES

1. C. Shan, S. Gong, and P. McOwan. "Facial expression recognition based on local binary patterns: A comprehensive study" *Image and vision Computing*, 27(6):803–816, May 2009.
2. G. Guo and C. Dyer., "Learning from examples in the small sample case: face expression recognition" *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 35, no. 3, pp. 477-488, Jun. 2005.
3. Q. Zhang, Z. Liu, G. Quo, D. Terzopoulos, and H. Y. Shum, "Geometry-driven photorealistic facial expression synthesis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 1, pp. 48-60, Feb. 2006.
4. G. Lei, X. H. Li, J. L. Zhou and X. G. Gong, "Geometric feature based facial expression recognition using multiclass support vector machines," *IEEE International Conference on Granular Computing*, 2009, GRC '09, pp. 318-321, Aug. 2009.
5. J. F. Ye, Y. Z. Zhan and S. L. Song, "Facial expression features extraction based on Gabor wavelet transformation" 2004. *IEEE International Conference on System, Man and Cybernetics*, vol. 3, pp. 2215-2219, Oct. 2004.
6. Q. Y. Zhao, B. C. Pan, J. J. Pan and Y. Y. Tang, "Facial expression recognition based on fusion of Gabor and LBP features", 2008. *ICWAPR '08. International Conference on Wavelet Analysis and Pattern Recognition*, vol. 1, pp. 362-367, Aug. 2008.
7. A. Koutlas and D. Fotiadis, "A region based methodology for facial expression recognition," *In BIOSIGNALS*, vol. 2, pp. 218–223, 2008.
8. T. Ojala, M. Pietikainen and D. Harwood, "A comparative study of texture measures with classification based on feature distributions" *J. Pattern Recognition* vol. 29, No.1 pp. 51-59, 1996.
9. Timo Ahonen, Abdenour Hadid, and Matti Pietikainen, "Face Recognition with Local Binary Patterns," *M. Lecture Notes in Computer Science*, Vol. 3021, pp.469-474, May.2004.
10. T. Ahonen, A. Hadid, and M. Pietikinen, "Face recognition with local binary patterns," *in ECCV*, 2004, pp. 469-481.
11. A. Hadid, M. Pietikinen, and T. Ahonen, "A discriminative feature space for detecting and recognizing faces," *in IEEE CVPR*, June 2004, pp. 797-804.
12. T. Ojala, M. Pietikainen, T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (7) (2002) pp. 971-987.
13. M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba. "Coding facial expressions with gabor wavelets," *In FG '98: Proceedings of the 3rd. International Conference on Face & Gesture Recognition*, pages 200–205, 1998.

14. T. Ahonen, A. Hadid, M. Pietikainen, "Face recognition with local binary patterns," in: European Conference on Computer Vision(ECCV), 2004.
15. Chuan- Yu Chang, Yan-Chuang Huang,Chi-lu Yang, "Pesonilized facial expression Recognition in Color Image'. Fourth International Conference on Inovative Computing Information and Control,pp1164-1167.
16. Xiaoyi Feng, "Facial Expression Recognition Based on Local Binary Patterns and Coarse-to-Fine Classification" Proceedings of the Fourth International Conference on Computer and Information Technology (CIT'04).
17. Di Huang, Caifeng Shan, Mohsen Ardabilian, Yunhong Wang and Liming Chen, "Local Binary Patterns and Its Application to Facial Image Analysis: A Survey," IEEE Transaction on Systems, Man and Cybernetics-part C: Applications and Reviews, vol.41, no.6, pp. 765-781, Nov 2011.
18. X. Feng, M. Pietikainen and A. Hadid, "Facial Expression Recognition with Local Binary Patterns and Linear Programming" Pattern Recognition and Image Analysis, vol. 15, No. 2, 2005, pp.546-548.
19. Mandeep Kaur, Rajeev Vashisht and Nirvair Neeru, "Recognition of Facial Expressions with Principal Component Analysis and Singular Value Decomposition," International Journal of Computer Applications, vol. 9, no.12,pp. 36-40, Nov 2010.
20. M.Pantic and L.Rothkrantz, "Automatic Analysis of Facial Expressions: The State of the Art", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 22, 2000, pp. 1424-1445.
21. B.Fasel and J.Luetin, "Automatic Facial Expression Analysis: A Survey", Pattern Recognition, Vol.36, 2003, pp. 259-275.
22. W.Fellenz, J.Taylor, N.Tsapatsoulis, and S.Kollias, "Comparing Template-based, Feature-based and Supervised Classification of Facial Expression from Static Images", Computational Intelligence and Applications, 1999.
23. M.Lyons, J.Budynek, and S.Akamastu, "Automatic Classification of Single Facial Images", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.21, 1999, pp. 1357-1362.
24. I. Cohen, N. Sebe, A. Garg, L. Chen and T.S. Huang, "Facial expression recognition from video sequences: temporal and static modeling", Computer Vision and Image Understanding 91 (2003), pp.160-187.

A Comprehensive Survey on Human Facial Expression Detection

Archana Verma

*Deptt. of Computer Science & Engineering,
Dr. C. V. Raman University
Bilaspur, 495113, India*

archana.ver2k@gmail.com

Lokesh Kumar Sharma

*National Institute of Occupational Health
Ahmedabad, 380016, India*

lksharmain@gmail.com

Abstract

In the recent years recognition of Human's Facial Expression has been very active research area in computer vision. There have been several advances in the past few years in terms of face detection and tracking, feature extraction mechanisms and the techniques used for expression classification. This paper surveys some of the published work since 2001. The paper gives a time-line view of the advances made in this field, the applications of automatic face expression recognizers, the characteristics of an ideal system, the databases that have been used and the advances made in terms of their standardization and a detailed summary of the state of the art. The paper also discusses facial parameterization using FACS Action Units (AUs) and advances in face detection, tracking and feature extraction methods. It has the important role in the human-computer interaction (HCI) systems. There are multiple methods devised for facial feature extraction which helps in identifying face and facial expressions.

Keywords: Facial Expression, FACS, Fuzzy Inference System, Feature Extraction, HCI.

1. INTRODUCTION

Facial features and expressions are critical to everyday communication. Besides speaker recognition, face assists a number of cognitive tasks: for example, the shape and motion of lips forming visemes can contribute greatly to speech comprehension in a noisy environment. While intuition may imply otherwise, social psychology research has shown that conveying messages in meaningful conversations can be dominated by facial expressions, and not spoken words. This result has led to renewed interest in detecting and analyzing facial expressions in not just extreme situations, but also in everyday human-human discourse. A very important requirement for facial expression recognition is that all processes therein have to be performed without or with the least possible user intervention. This typically involves initial detection of face, extraction and tracking of relevant facial information, and facial expression classification. In this framework, actual implementation and integration details are enforced by the particular application. For example, if the application domain of the integrated system is behavioral science, real-time performance may not be an essential property of the system.

A Facial Expression is a visible manifestation of the affective state, cognitive activity, intention, personality and psychopathology of a person [1]. Facial Expression convey nonverbal communication cues in face-to-face inter actions. Ekman and Freisen[1] have produced FACS – Facial Action Coding System for describing visually distinguishable Facial movements [2] [32]. Using the FACS, Action Parameters is designated to each of the expressions which classify the Human Emotions [2]. Also, Mehrabian[2] indicated that the verbal part of a message contributes for 7% to the effect of the message; the vocal part contributes 38% while facial

expression contributes for 55% [3]. There are wide range of applications of Facial Expressions, some of which include image understanding, psychological studies, facial nerve grading in medicine [4], face image compression and synthetic face animation [5]. Previous studies on the automatic analysis of Facial Expressions have discussed the classification methods and the extraction methods [6] [7] [8]. Although, there are various studies on the applications of computers in areas related to our proposed intelligent sales systems in the form of e-commerce and m-commerce, personalized online product selection, and web-based shopping systems we have, so far, found limited conceptual or empirical studies in this field. This paper is conceptual in nature and therefore aimed at exploring and initiating the debate in this field.

2. RELATED WORK

It has been well researched that facial expressions do reflect cognitive behaviour and that individuals observe other's facial expressions and then use these to regulate their own behavior in social interactions. We do find some studies relating facial expression to some aspects of marketing but none on facial expression and shopping behavior. Howard and Gengler [4] found that favourable facial expressions do have a positive bias on consumer product attitude. Sirakaya and Sonmez [5] used facial expression to study the gender images in Government tourism brochures. They describe a computerized intelligent sales assistant that gives sales personnel the ability to allocate their time where it will produce the best results, both for the customer, and for the business. Derbaix [6] also used facial expression to investigate the effect of television advertisement on attitude towards advertisement and brand attitude. They investigated 8- to 11-year-olds' reactions to advertising communication with a specific focus on the construct validity of the scales used to measure the main variables studied. Yuasa et al. [7] developed a computer network negotiation support tool using facial expression to negotiate an agreement strategy between seller and buyer. They argued that, if players select a happy face, then there is a greater chance they will reach an agreement. Lee [8] developed computer software programs to implement an e-shopping authentication scheme that used facial recognition instead of user name and password to control access to the web site in a modern consumer e-shopping environment. Consumers would not need to remember their user name and password to get access to the web site. The computer, based upon the facial features of the user, would recognize faces and allow access to the web site to authorized users.

A recent study in the context of computer based tutoring systems has shed some light on the types and frequencies of facial expressions exhibited by students during interactions with a computer. These are the same as the results that were found in another recent study of human tutoring sessions in which it is suggested that the expressions displayed by students are not significantly affected by whether the tutor is human or artificial. However, much more complete studies would be required before firm conclusions could be reached on various aspects of the facial expressiveness of customers.

The outcome of this research has some useful applications in the field of marketing, especially in sales (online or physical) and the training of sales people. There is also an added advantage that, if sales people can quickly identify the potential customers from window shoppers with the intervention of the system, then they can spend more time on potential customers and convert these potential customers into buyers. It will save the sales people time and the customer will feel well attended to. This will also reduce the expenses on sales forces as the stores will employ less people as they can identify the potential customer and devote sufficient time on them to convert them into buyers instead of spending unreasonable time on window shoppers. Using the same technology, online shoppers can be directed to appropriate products which could result in increased sales.

3. DIFFERENT FACIAL EXPRESSIONS

Prior to the compilation of the FACS in 1977, most of the facial behaviour researchers were relying on the human observers who would observe the face of the subject and give their analysis. But such visual observations cannot be considered as an exact science since the

observers may not be reliable and accurate. Ekman et al. [1] questioned the validity of such observations by pointing out that the observer may be influenced by context. They may give more prominence to the voice rather than the face and furthermore, the observations made may not be the same across cultures; different cultural groups may have different interpretations. Following Figure 1 demonstrates the seven universal expressions of emotion. Each of these expressions is racially and culturally independent.



FIGURE 1: Some Sample Images from The JAFFE Database.

The limitations that the observers pose can be overcome by representing expressions and facial behaviors in terms of a fixed set of facial parameters. With such a framework in place, only these individual parameters have to be observed without considering the facial behaviour as a whole. Even though, since the early 1920s researchers were trying to measure facial expressions and develop a parameterized system, no consensus had emerged and the efforts were very desperate [9] [10]. To solve these problems, in 1978, Ekman and Friesen [1] published the Facial Action Coding System (FACS), which, 30 years later, is still the most widely used method available. Through observational and electromyography study of facial behavior, they determined how the contraction of each facial muscle, both singly and in unison with other muscles, changes the appearance of the face. These changes in the face and the underlying (one or more) muscles that caused these changes are called Action Units (AU). The FACS is made up of several such action units. Figure 2 and 3 illustrates some of these Action Units.

NEUTRAL	AU 1	AU 2	AU 4	AU 5
Eyes, brow, and cheek are relaxed.	Inner portion of the brows is raised.	Outer portion of the brows is raised.	Brows lowered and drawn together	Upper eyelids are raised.
AU 6	AU 7	AU 1+2	AU 1+4	AU 4+5
Checks are raised.	Lower eyelids are raised.	Inner and outer portions of the brows are raised.	Medial portion of the brows is raised and pulled together.	Brows lowered and drawn together and upper eyelids are raised.
AU 1+2+4	AU 1+2+5	AU 1+6	AU 6+7	AU 1+2+5+6+7
Brows are pulled together and upward.	Brows and upper eyelids are raised.	Inner portion of brows and cheeks are raised.	Lower eyelids checks are raised.	Brows, eyelids, and checks are raised.

FIGURE 2: Upper Face Action Units and its Combination.

NEUTRAL	AU 9	AU 10	AU 12	AU 20
Lips relaxed and closed.	The infraorbital triangle and center of the upper lip are pulled upwards. Nasal root wrinkling is present.	The infraorbital triangle is pushed upwards. Upper lip is raised. Causes angular bend in shape of upper lip. Nasal root wrinkle is absent.	Lip corners are pulled obliquely.	The lips and the lower portion of the nasolabial furrow are pulled back laterally. The mouth is elongated.
AU 15	AU 17	AU 25	AU 26	AU 27
The corners of the lips are pulled down.	The chin boss is pushed upwards.	Lips are relaxed and parted.	Lips are relaxed and parted; mandible is lowered.	Mouth stretched open and the mandible pulled downwards.
AU 23+24	AU 9+17	AU 9+25	AU 9+17+23+24	AU 10+17
Lips tightened, narrowed, and pressed together.				
AU 10+25	AU 10+15+17	AU 12+25	AU 12+26	AU 15+17
AU 17+23+24	AU 20+25			

FIGURE 3: Lower Face Action Units and its Combination.

For example:

- AU 1 is the action of raising the Inner Brow. It is caused by the Frontalis and Pars Medialis muscles,
- AU 2 is the action of raising the Outer Brow. It is caused by the Frontalis and Pars Lateralis muscles,
- AU 26 is the action of dropping the Jaw. It is caused by the Masetter, Temporal and Internal Pterygoid muscles, and so on [10]. However not all of the AUs are caused by facial muscles.

Some of such examples are:

- AU 19 is the action of 'Tongue Out',
- AU 33 is the action of 'Cheek Blow',
- AU 66 is the action of 'Cross□Eye', and so on. The Face can be divided into Upper face [11] and Lower Face Action units [12] and the subsequent expressions are also identified. The Figures shows some of the combined action units.

4. STRUCTURE OF FACIAL ACTION

According to the manual of FACS [9] the facial expression analysis system estimates the measurement of actions and also classifies the actions. Various steps are involved in Facial Expression analysis i.e. 1. Face Acquisition 2. Facial Expression Extraction 3. Expression Recognition. Following Figure 4 shows the basic structure of facial expression analysis system.

Faces are detected from the input images or image sequences in the first step which is Face Acquisition phase. It can detect faces from input image or detect face form image sequence like detect face in the first frame and track the face in the remaining frames. After the face is located the facial features are extracted to identify the facial expression. Facial expression can be classified into two types namely Geometric or Intransient features and Appearance Features or Transient Features.

Geometric or Intransient Features: The features that are always present in the face but may be deformed due to any kind of facial expression.eg)Eyes, Eyebrows, Mouth, Tissue Textures, Nose. The facial components or facial feature points are extracted to form a feature vector that represents the face geometry.

Appearance or transient Features: The features that appear temporarily in the face during any kind of Facial Expression. Eg) Different kinds of wrinkles, bulges, forefront, regions surrounding the mouth and eyes. With appearance-based method, image filters such as Gabor wavelets [14] are applied to either the whole-face or specific regions in a face image to extract a feature vector. Facial Expression Recognition is the last step in facial expression analysis where the extracted features are recognized based on the action units. The Recognizer identifies not only the basic emotions like anger,happy,surprise,sad[13] but also identifies the expression caused due to pain[14],temporal dynamics [15] , Intensity of Expression[16],Spontaneous Expression [17].

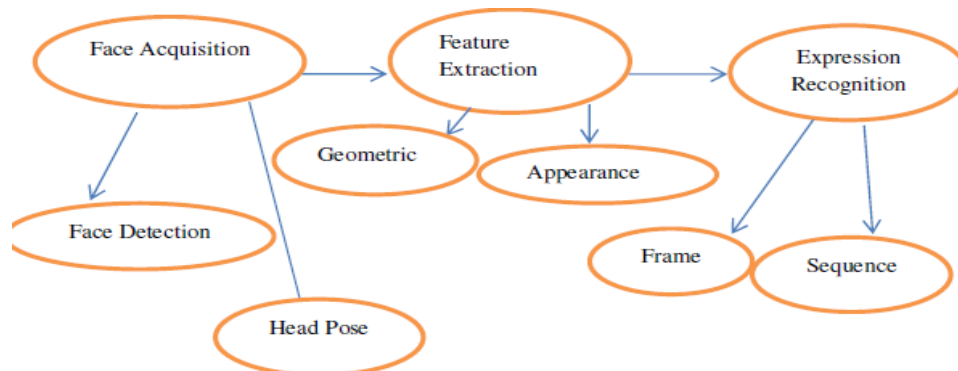


FIGURE 4: Basic Structure of Facial Expression analysis.

4.1 Face Acquisition

The two basic face acquisition methods are to detect faces both in frontal view images and near frontal view images. To detect the faces, two methods are used namely face detection and head pose estimation.

4.1.1 Face Detection

Many detection methods have been employed to detect faces [18] [19] [20] [21] [22]. Some of the previous face detection methods since 2003 are summarized here. Yeasin et al. [9] had used robust and automated face detection to segment the face region which was based on the work of Rowley [9]. Bartlett et al. [3] had designed boosting techniques for face detection in a generative framework based on their own work [17][18]. Tong et al. [10] had drawn geometry of face by encoding the edges of the face in the graph based on the work of Wiskott [10]. Kotsia and Patras [13] had employed convolution neural network for detecting the face and the classification is performed using a rule based algorithm [19].

4.1.2 Head Pose Estimation

To handle the out-of-plane head motion, head pose estimation can be employed. The methods for estimating head pose can be classified as 3D model-based methods [20] [21] and 2D image-based methods [28]. In 3D Model based method Bartlett used a canonical wire-mesh face model to estimate face geometry and 3D pose from hand-labeled feature points. In 2D image based method to handle the full range of head motion for expression analysis, Tian et al. [11] detected the head instead of the face. The head is identified using the smoothed silhouette of the foreground object as a segment using background subtraction and computing the negative curvature minima (NCM) points of the silhouette.

4.2 Facial Expression Extraction

Feature extraction can be viewed as a pre-processing step which removes distracting variance from a dataset, so that downstream classifiers or regression estimators perform better. The area where feature extraction ends and classification, or regression, begins is necessarily murky: an ideal feature extractor would simply map the data to its class labels, for the classification task. On the other hand, a character recognition neural net can take minimally pre-processed pixel values as input, in which case feature extraction is an inseparable part of the classification process. The extraction is basically based on the type of features, Geometric Features and Appearance Features. The two basic concepts employed for extracting features are based on identifying facial deformation and facial motion. The deformation based features recognize the Action Units, and the classifier is trained to differentiate human emotional states based on identified Action Units.

Methods		Geometric Features	Appearance Features
Deformation Extraction	Image Based	Gabor Filter[28]	Local Gabor Filter Bank [14], Fisher's Linear Decomposition, Singular value Decomposition [21] [22]
	Model Based	Point Distribution Model [20]	Feature point Tracking [25]
Motion Extraction	Frame Based	Active Contour (snake) [18]	Gabor Filter Bank [26]
	Sequence Based	PCA [23] [24], Gabor Filter Bank & AdaBoost [14]	Haar like feature [17], Multimodal facial feature Tracking [24], Candid Grid Node [16].

TABLE 1: Facial Expression Extraction Methods.

The deformation kind of extraction is applied to images and to image sequences [19]. The motion based features exploit the temporal correlation of facial expressions to identify variations within a probabilistic framework [15]. Image based models extract features from images, or reduced dimensional facial components [14]. Model based features are usually shape or texture models that fit human faces. The output of the feature extractor stage must contain separable and classifiable vectors. Active appearance models [22] and point distribution models [23] are used to fit on the shapes of interest. These shapes constitute the feature vectors. The expression

extraction methods are widely classified under two kinds namely deformation extraction and motion extraction. As for motion extraction techniques, some commonly used methods are dense optical flow [14], feature point tracking [5], and difference images [16]. The Various techniques under facial expression extraction methods are tabulated in the table.

4.2.1 Geometric Feature Extraction

Geometric Extraction is to detect and track changes of facial components in near frontal face images. Tian et al. [11] developed multi-state models to extract the geometric facial features. A three-state lip model describes the lip state: open, closed, tightly closed. A two-state model (open or closed) is used for each of the eyes. Each brow and cheek has a one-state model. Some appearance features, such as nasolabial furrows and crows-feet wrinkles (Figure 5), are represented.

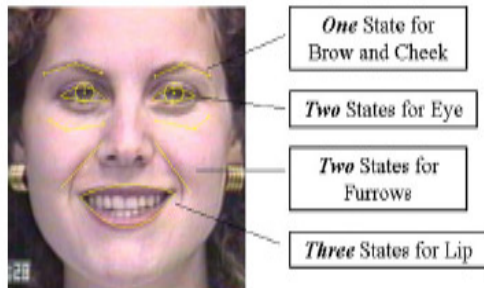


FIGURE 5: Geometric Feature Extraction.

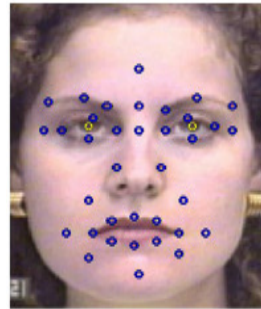


FIGURE 6: Appearance Feature Extraction.

Model Based: Automatic Active Appearance Model (AAM) mapping can be employed to reduce the manual preprocessing of the geometric feature initialization. Xiao et al. [11] performed the 3D head tracking to handle large out-of plane head motion and track no rigid features. Once the head pose is recovered, the face region is stabilized by transforming the image to a common orientation for expression recognition [22].

Image Sequence: Given an image sequence, the region of the face and approximate location of individual face features are detected automatically in the initial frame. The contours of the face features and components then are adjusted manually in the initial frame. After the initialization, all face feature changes are automatically detected and tracked in the image sequence. The system groups 15 parameters for the upper face [11] and 9 parameters for the lower face [12], which describe shape, motion, and state of face components and furrows. To remove the effects of variation in planar head motion and scale between image sequences in face size, all parameters are computed as ratios of their current values to that in the reference frame.

4.2.2 Appearance Feature Extraction

Gabor wavelets are widely used to extract the facial appearance changes as a set of multiscale and multi-orientation coefficients [14][20][28]. The Gabor filter may be applied to specific locations on a face or to the whole face image [16] [21]. There are two types of features to recognize expressions, the geometric positions of 34 fiducial points on a face and 612 Gabor wavelet coefficients extracted from the face image at these 34 fiducial points.

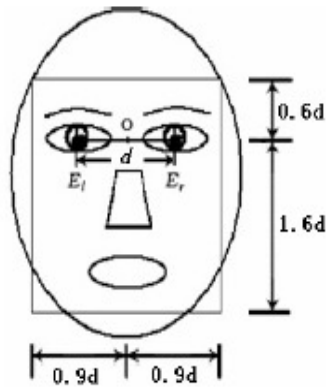


FIGURE 7: Facial Model.

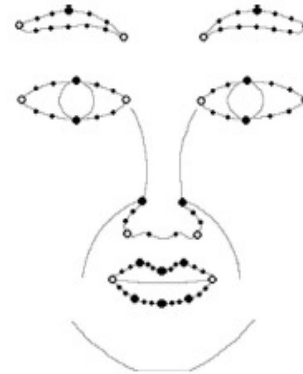


FIGURE 8: Feature Points in The Facial Model.

Fiducial points marked by circles (global) and big black dots (local), and contour points marked by small black dots. Image Sequence: Techniques like Haar-like Feature, Facial Feature tracking are used to identify the facial features that produces the expressions. A multi-modal tracking [24] approach is required to enable the state switching of facial components during the feature tracking process. Twenty-six fiducial points and 56 contour points are used in the facial model. Using the Facial model the fiducial points are marked for an image sequence using feature tracking method. The marked features in an image sequence are shown in the figure 8.



FIGURE 9: Feature Tracking in Image Sequence.

4.3.1 Frame-Based Expression Recognition

Frame-based expression recognition does not use temporal information for the input images. It uses the information of current input image with/without a reference frame. The input image can be a static image or a frame of a sequence that is treated independently. Several methods can be found in the literature for facial expression recognition such as neural networks [5][6], support vector machines [20], linear discriminant analysis [21], and rule-based classifiers [17].

4.3.2 The Sequence-Based Recognition

This method uses the temporal information of the sequences to recognize the expressions of one or more frames. To use the temporal information, the techniques such as HMM [29], recurrent neural networks, and rule-based classifier [17] were employed in facial expression analysis. The article by Yunfeng Zhu, Fernando De la Torre, Jeffrey F. Cohn [12] have made comparative studies for FACS AU recognition in spontaneously occurring behaviour by using the same RU-FACS database. Several systems had tried to recognize AUs or expression in spontaneously occurring behaviour [17].

The Expression recognition classifiers are listed below

Technique	Title	Year	Author (s)	Method	Database
Frame based	A New Facial Expression Recognition Method based on local Gabor filter bank and PCA plus LDA	2005	Deng et al. [13]	PCA plus LDA	JAFFE
	Automatic Facial Expression Recognition using facial animation parameters and multi-stream HMMS	2006	Aleksic et al. [14]	Multi-stream Hidden Markov Models	Cohn Kanade
	A Region Based methodology for facial expression recognition	2006	Koutlas and Fotiadis [15]	Neural Networks	JAFFE
	Automatic Recognition of Facial Actions in Spontaneous Expressions	2006	Bartlett et al. [16]	SVM & ADABOOST	RU-FACS
	A Facial Expression Classification System Integrating Canny, Principal Component Analysis and Artificial Neural Network	2001	Thai et al. [17]	PCA & ANN	JAFFE
Sequence Based	Recognition of Facial Expressions and Measurement of Levels of Interest From Video	2006	Yeasin et al. [18]	Hidden Markov Model	Cohn Kanade
	Boosting encoded dynamic features for facial expression recognition	2009	Yang et al. [19]	Adaboost	Cohn Kanade
	Recognizing Facial Expression: Machine Learning and Application to Spontaneous Behaviour[6]	2005	Bartlett et al. [20]	Adaboost & SVM	Cohn Kanade
	Robust facial feature tracking under varying face pose and facial expression	2007	Tong et al. [21]	Gabor filters & Switching Hypothesis Measurement	FRGC 1.0
	Facial Expression Recognition in Image Sequences Using Geometric Deformation Features and Support Vector Machines	2007	Kotsia and Patras [22]	Candid Grid Tracking, SVM, FAU	Cohn Kanade

	Dynamics of facial expression extracted automatically from video	2006	Littlewort et al. [4]	SVM	Cohn Kanade
	A Classifier Model based on the Features Quantitative Analysis for Facial Expression Recognition	2011	Jamshidnezhad and Nordin [22]	Fuzzy rule, Genetic Algorithm	Cohn Kanade

TABLE 2: Expression Recognition Methods.

6. CONCLUSION

The objective of this paper is to show a survey on the structure of analyzing the facial expression. The steps involved in expression analysis like face acquisition, feature extraction and expression classification had been discussed. Each step is discussed with the approaches and methods that can be applied to attain the required goal. Facial expression recognition is the key to next generation human-computer interaction (HCI) systems. We have chosen a more integrated approach as compared to most of the general applications of FACS. Extracted facial features are used in a collective manner to find out ultimate facial expression.

7. REFERENCES

1. P. Ekman and W. V. Friesen. "Facial Action Coding System", Consulting Psychologists Press Inc., 577 College Avenue, Palo Alto, California 94306, 1978.
2. Mehrabian, A., "Silent Messages", Wadsworth, Belmont, California, 1971.
3. M. S. Barlett et al. "Recognizing facial expression: machine learning and application to spontaneous behavior", IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 2005, pp. 568-573.
4. G. Littlewort et al. "Dynamics of facial expression extracted automatically from video", Image and Vision Computing Vol. 24, pp. 615-625, 2006.
5. Sirakaya E. and Sonmez S., "Gender Images in State Tourism Brouchures: An Overlooked Area in Socially Responsible Tourism Marketing", Journal of Travel Research, Vol. 38 (May): 353-362, 2000.
6. Derbaix, C. M. "The Impact of Affective Reactions on Attitudes toward the Advertisement and the Brand: A Step toward Ecological Validity", Journal of Marketing Research, Vol. 32 (November): 470-479, 1995.
7. Yuasa, M., Yasumura, Y. and Nitta, K., "A Negotiation Support Tool Using Emotional Factors", IEEE, Vol. 5: 2906-2911, 2001.
8. Lee, R.S.T., "iJADE Authenticator- An Intelligent Multiagent based Facial Authentication System", International Journal of Pattern Recognition and Artificial Intelligence, Vol.16, No. 4: 481-500, 2002.
9. Rowley, H., Baluja, S., and Kanade, T. 1998. Neural network-based face detection. IEEE Patt. Anal. Mach. Intell., 20:22-38.

10. L. Wiskott, J. M. Fellous, N. Kruger and C. von der Malsburg, Face recognition by elastic bunch graph matching, *IEEE Trans. Patt. Anal. Mach. Intell.* 19 (1997) 775{779. April 29, 2004 13:49 WSPC/115-IJPRAI 00322.
11. J. Xiao, S. Baker, I. Matthews, and T. Kanade. Real-time combined 2d+3d active appearance models. In *CVPR*, volume 2, pages 535 – 542, June 2004.
12. Yunfeng Zhu, Fernando De la Torre, Jeffrey F. Cohn, Associate Member, IEEE, and Yu-Jin Zhang, Senior Member, IEEE "Dynamic Cascades with Bidirectional Bootstrapping for Action Unit Detection in Spontaneous Facial Behavior", *Journal of LATEX Class Files*, October 2010
13. H. Deng, L. Jin, L. Zhen, and J. Huang. "A New Facial Expression Recognition Method based on Local Gabor Filter Bank and PCA plus LDA", *International Journal of Information Technology*, Vol. 11, No. 11, pp. 86-96, 2005.
14. P. Aleksic and A. Katsaggelos. "Automatic Facial Expression Recognition using Facial Animation Parameters and Multi-Stream HMMs", *IEEE Tran. on Information Forensics and Security*, Vol. 1, Issue 1, pp. 3-11, Mar. 2006.
15. A. C. Koutlas and D. I. Fotiadis. "A Region Based Methodology for facial expression recognition", *IEEE Int. Conf. on Systems, Man and Cybernetics*, 2008, pp. 662 - 666.
16. M. Bartlett, G. C. Littlewort , M. Frank, C Lainscsek, I. Fasel, J Movellan. "Automatic Recognition of facial actions in spontaneous expressions", *Journal of Multimedia*, Vol. 1, No.6, Sept. 2006.
17. L. Thai, N. Nguyen and T. Son. "A Facial Expression Classification System Integrating Canny, Principal Component Analysis and Artificial Neural Network" *International Journal of Machine Learning and Computing*, Vol. 1, No. 4, pp. 388-393, Oct. 2011.
18. M. Yeasin, B. Bulot, and R. Sharma. "Recognition of Facial Expressions and Measurement of Levels of Interest from video", *IEEE Transactions on Multimedia*, Vol.8, No.3, pp. 500-508 Jun 2006.
19. P Yang, Q. Liu, D. Metaxas. "Boosting Encoded dynamic features for facial Expression recognition", *Pattern Recognition Letters*, Vol. 30, pp. 132-139, Jan 2009.
20. M. S. Barlett et al. "Recognizing facial expression: machine learning and application to spontaneous behavior", *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, 2005, pp. 568-573.
21. Y. Tong, Y. Wang, Z. Zhu, J. Qiang. "Robust Facial Feature Tracking under varying face pose and facial expression", *Pattern Recognition*, Vol. 40, pp. 3195-3208, 2007.
22. I. Kotsia and I. Patras. "Facial Expression Recognition in Image Sequences using Geometric Deformation Features and SVM", *IEEE Transactions on Image Processing* Vol. 16, No.1, pp. 172-187, Jan 2007.
23. A. Jamshidnezhad, J. Nordin. "A Classifier Model based on the Features Quantitative Analysis for Facial Expression Recognition", *Proceeding of the International Conference on Advanced Science, Engineering and Information Technology*, 2011.
24. S. Lucey, A. Ashraf, and J. Cohn. "Investigating Spontaneous Facial Action Recognition through AAM Representations of the Face," *Face Recognition*, K. Delac and M. Grgic, eds., I-Tech Education and Publishing, 2007, pp. 275-286.

25. G. Fanelli, A. Yao, P. Noel, J. Gall, and L. Gool. "Hough Forest-based Facial Expression Recognition from Video Sequences", International Workshop on Sign, Gesture and Activity (SGA'10), Springer LNCS 6553, 2012, pp. 195-206.
26. P. Sharma, "Feature Based Method for Human Facial Emotion Detection using optical Flow Based Analysis", International Journal of Research in Computer Science, Vol. 1, Issue 1, pp. 31-38, 2011.
27. S. Koelstra, M. Pantic and I. Patras. "A Dynamic Texture-Based Approach to Recognition of Facial Actions and Their Temporal Models", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 32, No. 11, pp. 1940-1954, Nov. 2010.
28. D. Arumugam and S. Purushothaman. "Emotion Classification using Facial Expression" International Journal of Advanced Computer Science and Applications, Vol. 2, No.7, pp. 92-98, 2011.
29. S. Bashyal and G. Venayagamoorthy. "Recognizing facial expressions using gabor wavelets and vector quantization", Engineering Application of Artificial Intelligence, Vol. 21, 2008.
30. R. S. Smith, T. Windeatt. "Facial Expression Detection using Filtered Local Binary Pattern Features with ECOC Classifiers and Platt Scaling", JMLR: Workshop and Conference Proceedings 11, 2010, pp.111-118.
31. J.Q. Liu, Q. Z. Fan. "Research of feature extraction method on Facial Expression change", Advanced Materials Research, Vol. 211-12, pp. 813-817, Feb. 2011.
32. M. Leon, J. Rothkrantz. "Facial Action Recognition for Facial Expression Analysis from static face Images", IEEE Transactions on System and Cybernetics, Vol. 34, No. 3, pp. 1449-1461, Jun. 2004.
33. C. Shan, S. Gong and P. W. McOwan. "Facial expression recognition based on Local Binary Patterns: A comprehensive study" Image and Vision Computing, Vol. 27, pp. 803-816. May 2009.

Texture features from Chaos Game Representation Images of Genomes

Vrinda V. Nair

Professor in ECE

College of Engineering Trivandrum

Thiruvananthapuram, Kerala, India

vrinda66nair@gmail.com

Nisha N. S.

B.Tech student, College of Engineering Trivandrum

Thiruvananthapuram, Kerala, India

nish.phy@gmail.com

Vidya S.

B.Tech student, College of Engineering Trivandrum

Thiruvananthapuram, Kerala, India

vidyavasavan@gmail.com

Y. S. Thushana

B.Tech student, College of Engineering Trivandrum

Thiruvananthapuram, Kerala, India

ysthushana14@gmail.com

Abstract

The proposed work investigates the effectiveness of coarse measures of the Chaos Game Representation (CGR) images in differentiating genomes of various organisms. Major work in this area is seen to focus on feature extraction using Frequency Chaos Game Representation (FCGR) matrices. Although it is biologically significant, FCGR matrix has an inherent error which is associated with the insufficient computing as well as the screen resolutions. Hence the CGR image is converted to a texture image and corresponding feature vectors extracted. Features such as the texture properties and the subsequent wavelet coefficients of the texture image are used. Our work suggests that texture features characterize genomes well further; their wavelet coefficients yield better distinguishing capabilities.

Keywords: Chaos Game Representation, Texture analysis, Wavelet decomposition, Support Vector Machines.

1. INTRODUCTION

Chaos Game Representation (CGR) has been used in genomics and proteomics for various applications. Major work focuses on using Frequency Chaos Game Representation (FCGR) matrix for analysis. FCGR values though biologically relevant, has an inherent drawback due to insufficient screen as well as computing resolutions. In this work, attempt is made to investigate the effectiveness of coarse features such as texture properties and wavelet decomposition matrix of the texture image obtained from the gray scale equivalent of the corresponding FCGR. The results show that the coarse features of CGR extracted from the image in the form of texture as well as their wavelet coefficients can characterize genomes effectively.

Chaos game representation (CGR), the method used in this paper, for feature extraction, constructs a 2D image of the sequence data, which offers a visual understanding of the structure of the sequence. The differences between the various categories of sequences are evident from their respective CGR images. The CGR can be mapped into a numeric matrix by obtaining a Frequency CGR (FCGR) [1], [2]. A combined technique for genome classification using one probabilistic technique and two machine learning techniques based on FCGR features was

reported [3]. Hurst CGR a method using Hurst exponent to extract features from CGR is presented in [4].

This work maps the CGR image into its texture equivalent and corresponding properties are taken as features. The wavelet decomposition matrix of the texture image is also used as a feature vector for discriminating genome sequences.

2. MATERIALS AND METHODS

2.1 Dataset

Mitochondrial genomes are considered here. They are the sites of aerobic respiration, and are the major energy production center in eukaryotes. The low mutation rate in metazoan mitochondrial genome sequence makes these genomes useful for scientists assessing genetic relationships of individuals or groups within a species and for the study of evolutionary relationships [5]. Mitochondrial genomes were downloaded from the NCBI Organelle database [5]. Table 1 shows the data used for classification. The number of organisms shown is as listed in NCBI on 01/12/2012

Table 1: Dataset Used for Classification.

Serial number	Name of category	Number of organisms
1	Acoelomata	39
2	Cnidaria	48
3	Fungi	102
4	Plant	63
5	Porifera	44
6	Protostomia	582
7	Pseudocoelomata	63
8	Vertebrata	1729
	Total	2670

2.2 Methodology

2.2.1 Chaos Game Representation. The scope of CGRs as useful signature images of bio-sequences such as DNA has been investigated since early 1990s. CGR of genome sequences was first proposed by H. Joel Jeffrey [6]. To derive a chaos game representation of a genome, a square is first drawn to any desired scale and corners marked A, T, G and C. The first point is plotted halfway between the center of the square and the corner corresponding to the first nucleotide of the sequence, and successive points are plotted halfway between the previous point, and the corner corresponding to the base of each successive nucleotide. Mathematically, co-ordinates of the successive points in the chaos game representation of a DNA sequence is described by an iterated function system defined in Eq. 1 and Eq. 2

$$X_i = 0.5(X_{i-1} + g_{ix}) \quad (1)$$

$$Y_i = 0.5(Y_{i-1} + g_{iy}) \quad (2)$$

g_{ix} and g_{iy} are the X and Y co-ordinates respectively of the corners corresponding to the nucleotide at position i in the sequence [7]. The CGR of a random sequence gives a uniformly filled square. The CGR of DNA sequences plotted for various species gives images illustrating the non-randomness of genome sequences, which indeed means that the sequence has a structure, indirectly captured by the signature image. Features of CGRs include marked double scoops, diagonals, varying vertical intensities, absence of diagonals etc. signifying corresponding sequence characteristics. The CGR is thus found to be unique for every species. Hence CGR of

genomic sequences are expected to furnish features of discriminative nature which could subsequently be presented to classifiers.

2.2.2 Texture Analysis. Texture analysis refers to the branch of imaging science that is concerned with the description of characteristic image properties by textural features. Texture analysis provides unique information on the texture, or spatial variation of pixel [8]. In texture analysis, a pixel occurrence probability matrix and a gray co-matrix, both obtained from the grayscale image is considered. The gray scale image is obtained by converting the FCGR matrix values into equivalent gray values. A feature vector is formulated which is an eight element matrix, which are actually eight properties of the image. Out of these eight properties, four - variance, skewness, kurtosis and entropy are obtained from pixel occurrence probability matrix and the other four - contrast, correlation, energy and homogeneity are obtained from gray co-matrix. Thus a feature vector having eight elements corresponding to the eight properties of the texture image characterizing the organism is extracted. Fig. 1 gives the CGR image of NC_000928 Echinococcus multilocularis and Fig. 2 the corresponding gray image,

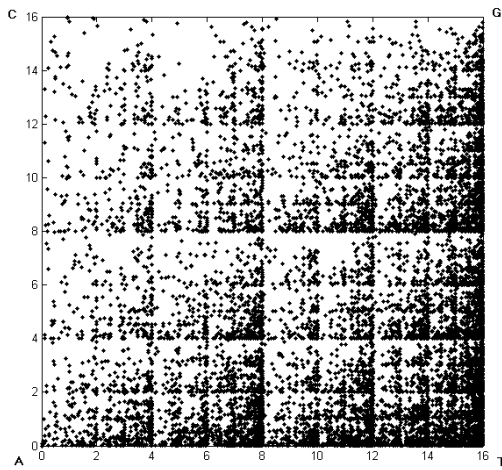


FIGURE 1: CGR Image of NC_000928 Echinococcus Multilocularis.

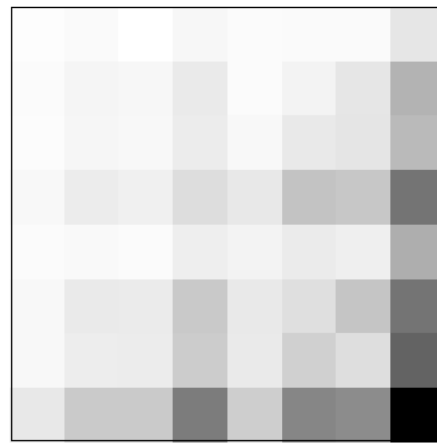


FIGURE 2: Grayscale Equivalent of CGR Image Shown in Figure 1.

2.2.3 Pixel occurrence probability matrix. First-order texture analysis measures use the image histogram, or pixel occurrence probability, to calculate texture. The main advantage of this approach is its simplicity through the use of standard descriptors (e.g. mean and variance) to characterize the data [9].

Assume the image is a function $f(x,y)$ of two space variables x and y , $x=0,1...L-1$ and $y=0, 1... M-1$. The function $f(x,y)$ can take discrete values $i = 0, 1....G-1$, where G is the total number of number of intensity levels in the image. The intensity-level histogram is a function showing (for each intensity level) the number of pixels in the whole image. The histogram contains the first-order statistical information about the image (or its fragment). Dividing the number of pixels having a given intensity value by the total number of pixels in the image gives the approximate probability density of occurrence of the intensity levels [10]. If $N(i)$ is the number of pixels with intensity i and M is the total number of pixels in an image, it follows that the histogram, or pixel occurrence probability, is given by

$$P(i)=N(i)/M \quad (3)$$

Out of several properties of the texture image, 4 were selected which showed maximum variation in values for different organisms under different classes. The four properties are variance, skewness, kurtosis, and entropy. For a random variable X with mean μ and standard deviation σ and expectation value E , the different properties are:

$$\text{Variance: } \text{var}(X) = E[(X - \mu)^2] \quad (4)$$

$$\text{Skewness: } \frac{E(X - \mu)^3}{\sigma^3} \quad (5)$$

$$\text{Kurtosis: } \frac{E(X - \mu)^4}{\sigma^4} \quad (6)$$

$$\text{Entropy: } - \text{sum}(p. * \log 2(p)) \quad (7)$$

p is the histogram counts returned from histogram image.

The variance is a measure of the amount of variation of the values of that variable from its expected value or mean. The skewness is a measure of asymmetry of the data around the sample mean. The skewness is zero if the histogram is symmetrical about the mean, and is otherwise either positive or negative depending whether it has been skewed above or below the mean. The kurtosis is a measure of flatness of the histogram. For a normal distribution the kurtosis is three and for other cases it will be greater than or less than three. Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image [11]. These four properties constitute the four elements of the feature vector obtained from the texture analysis of the image.

2.2.4 Graycomatrix

The different properties of the graycomatrix are known as the graycoprops. There are four properties for this matrix. The graycomatrix creates a gray-level co-occurrence matrix (GLCM) from an image. Graycomatrix creates the GLCM by calculating how often a pixel with gray-level (grayscale intensity) value i occurs horizontally adjacent to a pixel with the value j . Each element (i, j) in the GLCM specifies, the number of times that the pixel with value i occurred horizontally adjacent to a pixel with value j . The graycomatrix calculates the GLCM from a scaled version of the image. By default, if 'I' is a binary image, graycomatrix scales the image to two gray-levels. If 'I' is an intensity image, graycomatrix scales the image to eight gray-levels [10].

Graycoprops normalizes the gray-level co-occurrence matrix (GLCM) so that the sum of its elements is equal to 1. Each element (r,c) in the normalized GLCM is the joint probability occurrence of pixel pairs with a defined spatial relationship having gray level values r and c in the image. Graycoprops uses the normalized GLCM to calculate properties. The four properties are:

Contrast: It returns a measure of the intensity contrast between a pixel and its neighbor over the whole image. Contrast is 0 for a constant image.

Correlation: It returns a measure of how correlated a pixel is to its neighbor over the whole image. Its range is between -1 and +1. Correlation is 1 or -1 for a perfectly positively or negatively correlated image. Correlation is 'NaN' (not a number) for a constant image.

Energy: It returns the sum of squared elements in the GLCM. Its range is between 0 and 1. Energy is 1 for a constant image.

Homogeneity: It returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Its range is between 0 and 1. Homogeneity is 1 for a diagonal GLCM [10], [12].

2.2.5 Wavelet decomposition. The wavelet decomposition of a signal $f(x)$ is performed by a convolution of the signal with a family of basis functions. In the case of two-dimensional images, the wavelet decomposition is obtained with separable filtering along the rows and along the columns of an image. The wavelet analysis can thus be interpreted as image decomposition in a set of independent, spatially oriented frequency channels. The HH sub image represents diagonal

details (high frequencies in both directions – the corners), HL gives horizontal high frequencies (vertical edges), LH gives vertical high frequencies (horizontal edges), and the image LL corresponds to the lowest frequencies. At the subsequent scale of analysis, the image LL undergoes the decomposition using the same g and h filters, having always the lowest frequency component located in the upper left corner of the image [13]. In the case of a 3-scale analysis, 10 frequency channels can be identified. The size of the wavelet representation is the same as the size of the original image. As there is a choice of particular wavelet function for image analysis, symmetric wavelet functions appear superior to non-symmetric one, which is attributed to the linear-property of symmetric filters.

2.2.6 Support Vector Machines. Support Vector Machine classifier. Support Vector Machine was introduced to solve dichotomic classification problems [14] & [15]. Given a training set in a vector space, SVMs find the best decision hyper plane that separates two classes. The quality of a decision hyper plane is determined by the distance between two hyper planes defined by support vectors. The best decision hyper plane is the one that maximizes this margin. SVM extends its applicability on the linearly non-separable data sets by either using soft margin hyper planes or by mapping the original data vectors into a higher dimensional space in which the data points are linearly separable. There are several typical kernel functions. In this work, Support Vector Machine with Radial Basis kernel function and Polynomial kernel functions are used.

3. RESULTS AND DISCUSSION

The work aims to investigate the quality of features derived from the texture analysis and wavelet decomposition of a grey scale image of CGR (Chaos Game Representation) plot of each organism, evaluated through classification. The data set used was mitochondrial DNA sequences. The mitochondrial DNA sequences (DNA in mitochondria of a cell) of 2670 eukaryotic organisms belonging to eight categories of taxonomical hierarchy were obtained from National Centre for Biotechnology Information (NCBI) organelle database.

The total number of organisms in each class is first divided in 1:1 ratio to get two data sets as test and train. The Chaos game representation (CGR) of each sequence is obtained. Subsequently the FCGR matrix is computed and the corresponding gray scale images plotted. The feature vector, which corresponds to different properties of the grey image, is obtained for the different organisms in eight classes. These feature vectors are given as the input to the SVM classifier which is trained using the training set and tested using the test set. Using wavelets, 3 levels of decompositions were considered.

Previous works report using FCGR matrix elements as features for analysis of genomes [1], [2], [3], [7]. For huge sequences, since the screen resolution and computing resolution is limited, there will be error while computing the FCGR matrix. Hence this method is a novel attempt to provide an alternative to FCGR in such cases where huge sequences are involved. The work proves that the texture features as well the wavelet coefficients could be potential elements representing features of genomes. Other image transform coefficients as well as other image features can also be subjected to investigation in future, which may prove to be better representatives of genome sequences.

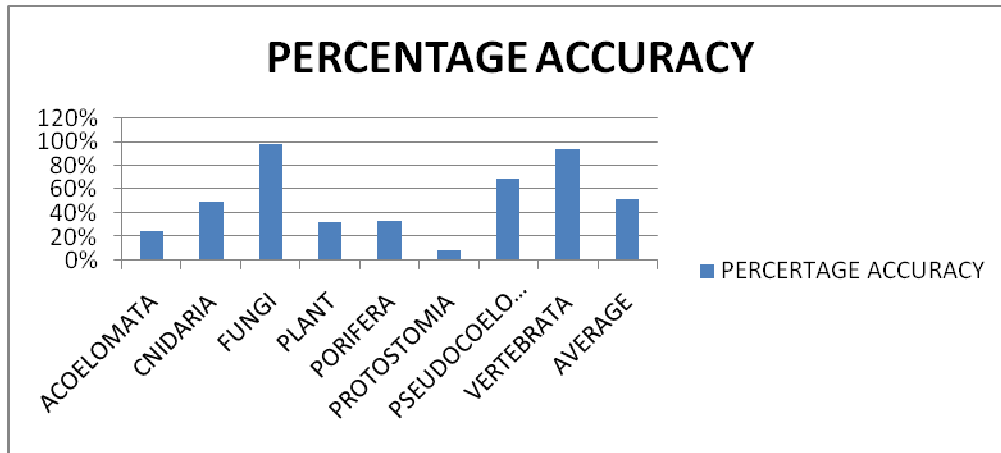


FIGURE 3: Percentage Accuracy of Classification Using Texture Analysis.

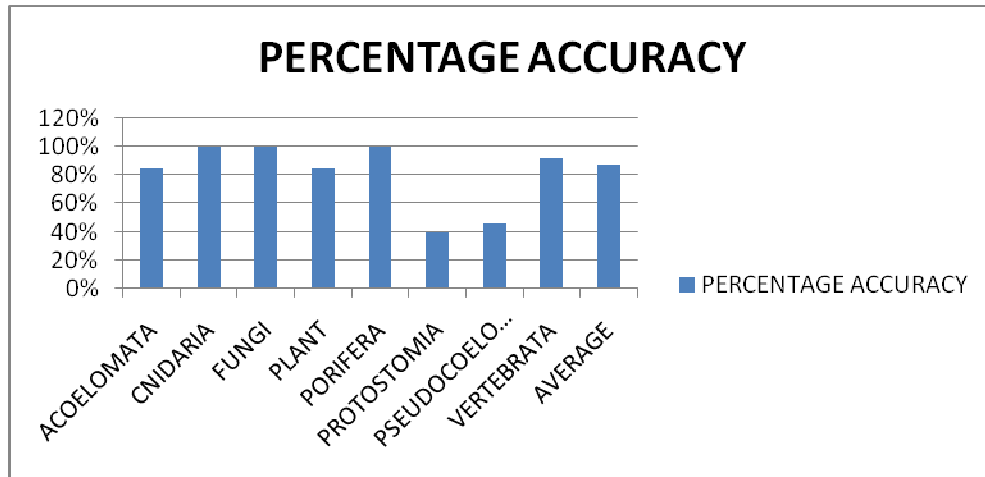


FIGURE 4: Percentage Accuracy of Classification Using 3 Level Wavelet Decomposition.

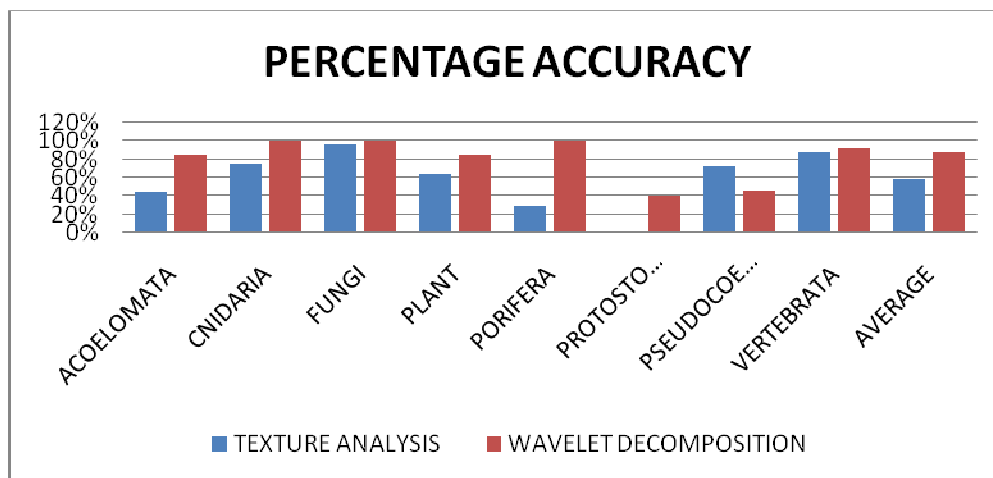


FIGURE 5: Texture Features and Wavelet Coefficient of The Texture Image – A Comparison of Feature Vectors Based on Classification Accuracy.

4. CONCLUSION

This work is an investigation into the quality of features derived from the texture analysis and wavelet decomposition of a grey scale image of CGR(Chaos Game Representation) plot of genomes evaluated through classification of organisms. The feature vectors were used to classify organisms with the help of an SVM classifier. The accuracy of classification stands testimony to the possibility of deriving feature vectors from the texture equivalent of CGR or more precisely – the FCGR matrix thus overcoming the inherent resolution error in FCGR matrices when considered quantitatively. It is thus concluded that coarse features such as texture and wavelet coefficients thereof characterise the genome sequences effectively.

5. REFERENCES

1. J. S Almeida, J. A. Carrico, A. Maretzek, P. A. Noble and M. Fletcher, "Analysis of genomic sequences by chaos game representation." *Bioinformatics*, vol. 17, (5), pp. 429–437, Jan. 2001.
2. P. J. Deschavanne, A. Giron, J. Vilain, G. Fagot and B. Fertil, "Genomic signature: characterization and classification of species assessed by chaos game representation of sequences." *Mol. Biol. Evol.*, vol 16, pp. 1391–1399, Jun. 1999.
3. V. V. Nair and A. S. Nair, "Combined classifier for unknown genome classification using chaos game representation features" in *Proc. International Symposium of Bio Computing*, 15-17 February 2010.
4. V. V.Nair, A. Mallya, B. Sebastian, I. Elizabeth, and A. S. Nair, *Hurst CGR (HCGR) – A Novel Feature Extraction Method from Chaos Game Representation of Genomes: ACC2011 Springer Proceedings*, June 2011.
5. Internet: <http://www.ncbi.nlm.nih.gov/Genomes/ORGANELLES/organelles.html>, [01/03/2012].
6. H. J. Jeffrey, "Chaos game representation of gene structure". *Nucleic Acids Res.*, vol. 18,8, pp. 2163–2170, Mar. 1990.
7. J. Joseph and R. Sasikumar, "Chaos game representation for comparison of whole genomes". *BMC Bioinformatics*, vol. 7, 243, May 2006.
8. A. Materka, M. Strzelecki, "Texture Analysis Methods – A Review.", University of Lodz, Institute of Electronics, COST B11 report, Brussels 1998.
9. W. H. Nailon, "Texture Analysis Methods for Medical Image Characterization", Department of Oncology Physics, Edinburgh Cancer Centre & School of Engineering, University of Edinburgh, United Kingdom.
10. Manual: MATLAB, R2010a, Image Analysis and Statistics, Texture Analysis, Graycomatrix, The MathWorks Inc.
11. D. Avola, L. Cinque and G. Placidi, "Medical Image Analysis Through A texture based Computer Aided Diagnosis Framework." *Int. J. Biomet. Bioinf.*, vol. 6, (5), pp. 144-152, Oct. 2012.
12. S. A. Angadi and M. M. Kodabagi, "A Texture Based Methodology for Text Region Extraction from Low Resolution Natural Scene Images." *Int. J Image Proc.*, vol. 3, (5), pp. 229-245, Nov. 2009.

13. K.P. Soman, K. I. Ramachandran and N. G. Resmi, Insight into Wavelets, PHI Learning Pvt. Ltd., 2010.
14. N. Cristianini and J.S. Taylor, Support vector machines and other kernel-based learning methods. Cambridge University Press, Cambridge , 2000.
15. V. N., Vapnik. The Nature of Statistical Learning Theory. Berlin: Springer-Verlag, 1995.

Diagnosis of Burn Images using Template Matching, k-Nearest Neighbor and Artificial Neural Network

Malini Suvarna

*Research Scholar,
Dept. of E&C,
Atria Institute of Technology, Bangalore*

maliniatria@gmail.com

Dr. Sivakumar

*Head Dept. of TC,
Dr. Ambedkar Institute of Technology, Bangalore*

sivabs2000@yahoo.co.uk

Dr. Kamal Kumar

*Plastic Surgeon,
Rajarajeshwari Medical College, Bangalore*

drkamalkumar25@yahoo.co.in

Dr. U C Niranjana

*Director of Research and Training,
Manipal Dot Net, Manipal*

ucniranjana@yahoo.com

ABSTRACT

The aim of this research is to develop an automated method of determining the severity of skin burn wounds. Towards achieving this aim, a database of skin burn images has been created by collecting images from hospitals, doctors and the Internet. The initial pre-processing involves contrast enhancement in lab color space by taking luminance component. Various pattern analysis or pattern classifier techniques viz. Template Matching (TM), k Nearest Neighbor Classifier (kNN) and Artificial Neural Network (ANN) have been applied on skin burn images and a performance comparison of the three techniques has been made. The help of dermatologists and plastic surgeons has been taken to label the images with skin burn grades and are used to train the classifiers. The algorithms are optimized on pre-labeled images, by fine-tuning the classifier parameters. During the course of research, of the three classifier methods used for classification of burn images it has been observed that the ANN technique reflected the best results. This has been inferred based on the comparative studies of the three methods. In the ANN method the classification of the image of burns has been found to be the nearest to the actual burns. The efficiency of the analysis and classification of the ANN technique has been of the order of 95% for Grade-1 burns, 97.5% for Grade-2 burns and 95% for Grade-3 burns. As compared to 55%, 72.5% and 70% for Grade1, Grade2, and Grade 3 burns respectively for the TM Method and 67.5%, 82.5% and 75% for kNN method. It is therefore felt that the ANN technique could be applied to analyze and classify the severity of burns. This burn analysis technique could be safely used in remote location where specialists' services are not readily available. The local doctors could use the analyzer and classify the grade of the burn with a good degree of accuracy and certainty. They could start preliminary treatment accordingly, prior to specialists' services. This would definitely go a long way in mitigating the pain and sufferings of the patients.

Keywords: ANN, TM, kNN.

1. INTRODUCTION

Medical Imaging is a boon given by science and technology to humanity. While on the one hand medical diagnostics has advanced in leaps and bounds, on the other hand advancement in image processing, pattern recognition and machine intelligence techniques has intensified medical imaging. With these advancements it is now possible to make diagnostics in a non-invasive manner. An easy and correct diagnostic facilitates appropriate treatment earlier and thereby enhance chances of full recovery, while reducing the pain and suffering of patients. Modern techniques are also cost effective.

For a successful evolution of a burn injury it is essential to initiate the correct first treatment [1]. To choose an adequate one, it is necessary to know the depth of the burn, and a correct visual assessment of burn depth highly relies on specialized dermatological expertise. As the cost of maintaining a burn unit is very high, it would be desirable to have an automatic system to give a first assessment in all the local medical centers, where there is a lack of specialists [2], [3]. Burn injury is one of the major accidents and life threatening causes in the modern world. Handling and management of burn victims is done in special wards in hospitals by specially trained personnel. Skin is the largest organ of our body and it gets damaged predominantly during burn accidents. Skin accounts for 15% of the total weight of an adult human being. The basic functions of the skin are protection, sensation, and temperature regulation, synthesis of vitamin D. The principal components of the human skin are epidermis and dermis. Epidermis is the outer thinner part of the skin, while the dermis is the inner thick layer of connective tissue made of elastic fibers. The World Health Organization demands that, at least, there must be one bed in a Burn unit for every 500000 inhabitants. So normally, one Burn Unit covers a large geographic extension [4]. If a burn patient appears in a medical center without Burn unit, a telephone communication is established between the local medical center and the closest hospital with Burn Unit, where the non-expert doctor describes subjectively the color, texture and status considered important for burn characterization. The result in many cases is the application of an incorrect first treatment (very important on the other hand for a correct evolution of the wound), or unnecessary displacements of the patient involving high sanitary cost and psychological trauma for the patient and family.

When a person meets with a burn accident the skin layers get affected. Doctors determine the degree of burns by examining which layer and organs are affected and suggest treatment. The work proposed here focuses on automating this process. A digital camera is used to capture the burn images of the patient and the software developed would analyze the image. Using the outcome of this research, severity of the injury can be estimated, the degree of burns can be calculated and the depth of the injured tissue can be quantified.

Skin Burns Images are collected from open source database of images and also from the burn ward of some hospitals. The images are digital and color. In order to assess the progress of healing of wound, a series of images of wound subjected to medication and palliative care are acquired. A data base of such images with substantial number was constructed to validate the outcome of this research. The general and clinical details of the subjects such as gender, age, ethnicity, treatment history, cause and type of burn and experts opinion is collected for each image.

General image processing and enhancement algorithms applicable to wound characterization is applied in this work on medical skin burn images. Usual flow of computer based processing begins with contrast enhancement, feature extraction from the red, green and blue components such as variance, mean, hue, saturation, intensity, texture, shape, and area. As there are no invasive techniques to assess the nature of the wound, computerized techniques are proposed as a cost effective solution.

The proposed automatic wound analyzer would be useful at remote locations where medical experts are not available. During accidents as the patient gets admitted to the hospital or while being transported in ambulance, the medical personnel attending the patient can send burn images immediately through online camera to the specialist. When the images reach the specialist, he can decide the line of treatment. He can diagnose whether surgical intervention is required or the patient can be treated by palliative methods. This will aid the patient recovery, it will also serve as a good teaching and research tool for students of health care. This would also serve as an aid in the follow up and further pathology for the burns. The proposed system is built and optimized such that it would use little computational resources, power and would compute faster.

The proposed algorithm software is first developed on the Matlab with the help of Image processing toolbox to enhance and pre-process the raw images [5, 6]. Proposed System is shown in Figure 1.1.

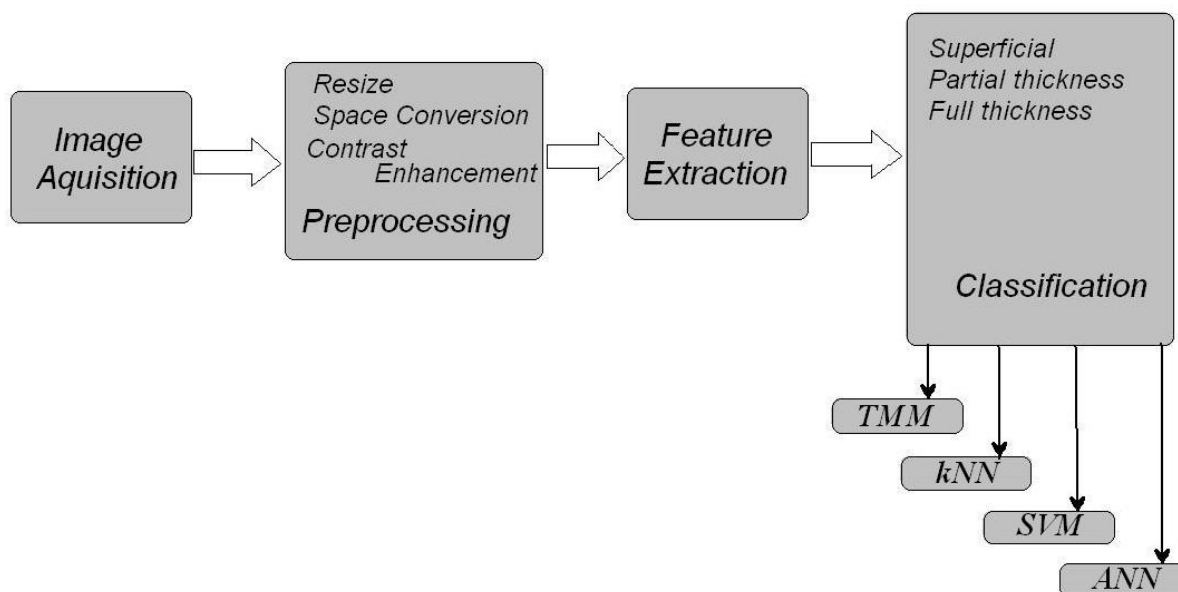


FIGURE 1.1: Proposed Burn Image Diagnosis System.

1.1 Literature Survey

Ongoing experimental research on skin burn images is very less. Thus, contribution to this field of research is not much. Contributions have been done particularly by the researchers from Spain, belonging to Biomedical Engineering Group of Seville University. Though the work is of significant importance, very little research has been done.

The following is the list of researchers who have contributed to this field of activity:

Gary L. Hansel et.al. [7]: This researcher developed clinical evolution of burn injuries using an optical reflectance technique.

Gary L. Hansel et.al.[9]: have proposed a method for the diagnosis of burns and pressure ulcers in the early stage using computerized image processing. These authors have diagnosed burns and pressure ulcers in the early stages, using color image processing, and have developed a method for quantifying the histological readings and have applied these readings to model wound formation. By making the color analysis and taking the hue factor of the wound, mild, moderate and severe types of the wounds are detected, using the time of injury as known variable.

Jean Phillipe.et.al. [16]: Have developed an algorithm to detect cancerous tissues from an image of microscopic section. Based on the shape and size of the cell, using mathematical morphology the tissues are classified into malignant or benign. Depending on the mathematical value obtained for each cell, the tissues are categorized into four groups.

Martin.et.al. [8]: They designed a new clinical instrument for evaluating burn depth. The imaging burn depth indicator produces a true color and a false color image of burn. The false color image consists of upto four colors, each of which indicates a distinct range of probability that the area of burn so colored will heal with in 21 days.

J K Bennet et.al. [12]: They have evaluated the burn depth by the use of radio active isotope.

Serrano et.al. [2]: The research scholars belonging to the Biomedical Engineering Group of Seville University have made a study to find the effectiveness of telemedicine for plastic surgery applications. They have considered burn images for tele-diagnosis by capturing the burn image in a digital camera and compressing it for transmission through a communication media

Begona Acha et.al. [11]: have proposed a method classifying burn into their depth, based on features extracted from color and texture characteristic of burn images. A fuzzy art map neural network and SVM classifier used for this work.

Begona Acha et.al. [10]: have proposed a method to separate burn skin from a normal skin in burn color images and to classify the according to the depth of the burn. In the classification part we take advantages of color information by clustering, with the vector quantization algorithm. The color centroids of small squares taken from the burnt segmented part of the images in the (V_1V_2) plane in two possible groups where V_1V_2 are the two chrominance component of the CIE lab representation.

Begona Acha) et.al. [4]: Here the author classifies burn into different grades (of depth) based on feature extracted from the color and texture characteristic of burn images addressed. A Fuzzy-ATMAP neural network and non-linear SVM with the different type of kernels have been compared.

2. CLASSIFICATION OF BURN INJURIES

Classification of the burn injury in this work, mainly depends on the color of the wound. Injury to the top layer of skin epidermis is called superficial burn. Injury to the second layer of skin dermis is called a partial thickness or dermal injury. An injury that extends down to the third layer subcutaneous tissue which includes fat is called a full thickness injury.

2.1 Superficial (Grade 1) Burn

In superficial burn epidermis layer of the skin gets affected. Best example is sun burn heal with in 5 to 7 days. Superficial burn is as shown in Figure: 2.1

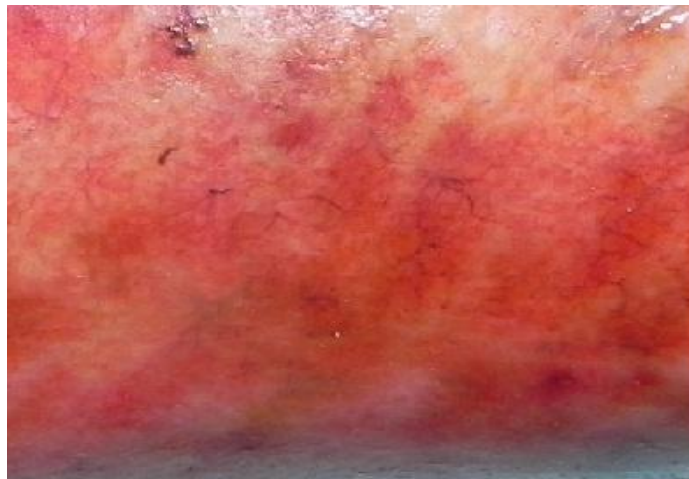


FIGURE 2.1: Superficial Dermal Burn.

2.2. Partial Thickness (Grade 2) Burn

In the Partial thickness burn dermis layer of the skin gets affected. Partial thickness burns usually leave scars. This will usually be treated with skin grafting Figure 2.2 shows a deep and large partial thickness burn.



FIGURE 2.2: Partial Thickness Burn.

2.3. Full Thickness (Grade 3) Burn

A full thickness burn destroys epidermis, dermis and subcutaneous layers of skin. Figure 2.3 shows a Full Thickness burn.



FIGURE 2.3: Full Thickness Burn.

3. DATABASE

Skin Burn Images of different grades are collected from hospital and Internet and scanned from biomedical books. Our database consists of a total of 120 images, 40 images from each grade of burn, as shown in Table 3.1. These images are labeled by the plastic surgeon. These images are used in the image classification. Once the algorithm is optimized on pre label images, it would be used to analyze non labeled images. Contrast enhancement and classification algorithm are applied to these images.

Type	No. of Images	Internet	Captured	Scanned from Books
Grade 1	40	8	27	5
Grade 2	40	7	28	5
Grade 3	40	4	32	4

Table 3.1: Database of Burn Images.

4. FEATURE EXTRACTION

Feature selection is very important while classifying the skin burn image into different grades. The selected features represent the characters of the images belonging to a particular category. Since the color of the skin burn images differs based on the depth of the wound, the color features of each image is extracted and used for the classifier training [1, 2].

In the research, initially the image is re-sized to 90*90 pixels, and then Red, Green and Blue (RGB) space is converted in to L*a*b* color space. Lab color space has 3 coordinates, one luminous and two chrominance V1 and V2. Luminous component is used for contrast enhancement. After contrast enhancement of the image, the V1 and V2 chrominance planes of the L*a*b* color space is selected for feature extraction. Further a 90*90 image is subdivided into 9*9 blocks and the features like mean and (2, 1)th coefficient of Discrete Cosine Transform (DCT) function is chosen to train classifiers. The two dimensional DCT equation is given below, where X(k1, k2) is the DCT and x(n1, n2) is the image

$$X(k_1 k_2) = \frac{4 \epsilon_{k_1} \epsilon_{k_2}}{N^2} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x(n_1, n_2) \cos\left(\frac{\pi(2n_1 + 1)k_1}{2N_1}\right) \cos\left(\frac{\pi(2n_2 + 1)k_2}{2N_2}\right)$$

$$\text{where } \begin{cases} \epsilon_k = \frac{1}{\sqrt{2}} & \text{for } k = 0 \\ 1 & \text{otherwise} \end{cases}$$

$$k = 0, 1, 2, \dots, N$$

5. TEMPLATE MATCHING METHOD

This is a very simple and straight forward method of classification. Images of various classes of burns are first classified and systematically stored as reference templates. The burn images are compared with the reference templates to classify the sample burn images. Where the sample images match the burn image well, the error is would be the least. The figure 5.1 given below illustrates the comparison of the sample images with a template. This type of comparison does not involve much computation. It is restricted to the number of samples to be compared and the number of templates available in the bank, proportionate to which the duration of time taken to compare and throw up a result would vary. Another limitation of the method is that when exact matches are not found the probability of mis-prediction increases. In this method the computational time increases as the number of sample increases [13].



FIGURE 5.1: Template Matching Method.

6. K-NEAREST NEIGHBOR CLASSIFIER

The possibility of an exact match between reference images and sample image not being found in the template method is addressed in this method. In this classification method a group of 'k' images from the reference images is selected which are the nearest match to the sample image. A label is assigned to the sample images based on the nearest 'k' image matching. The possibility of classification error where the image lies on the border of two classes is also overcome by this method where the labeling is based on average of 'k' images and not one reference image.

The steps involved in kNN classifier algorithm are:

1. A bank of labeled objects (burn images) is created. This bank serves as reference images for comparison. It is also called a "Training Set 'D'".
2. A distance or similarity metric that can be used to compute the closeness of sample objects is set up.
3. A value 'k' is generated which indicates the number of 'nearest' or 'closest' classes to the sample objects (images). The value 'k' varies depending on the intensity of the burn which is reflected in the image.
4. The 'k' nearest objects are used to determine the class of the target object or burn.

Consider a burn target object which needs to be classified. A set of images 'z' of the target object is compared with a training set of D images already set up. The algorithm compares nearness or closeness of the 'z' images with the set of D and computes k images which can be considered to be the nearest neighbors of the 'z' images of the target object. The value of 'k' images is generated by selecting images of 'D' based on their frequency in nearness to the 'z'.

The algorithm has a storage complexity of $O(n)$, where n is the number of training objects or images. Since the Euclidean distance of every test image needs to be computed with reference to the test objects the time complexity is also $O(n)$. Since classification models viz. a decision tree or separating hyper plane is not constructed there is no time lost for this activity as in most other classification methods. While other classification methods are costly proportionate to the model building stages, kNN classification method is different and at the same time comparatively inexpensive since classification steps are constant 'O'.

Considering a data matrix $m \times n$ where n reflects row vectors and data matrix of Y represented by $m \times n$, the euclidean distance between the vector x_s and y_t is defined as:

$$d_{st}^2 = (x_s - y_t) (x_s - y_t)'$$

Selection of k value is very important in kNN classifier, this is explained in the Figure 6.1.

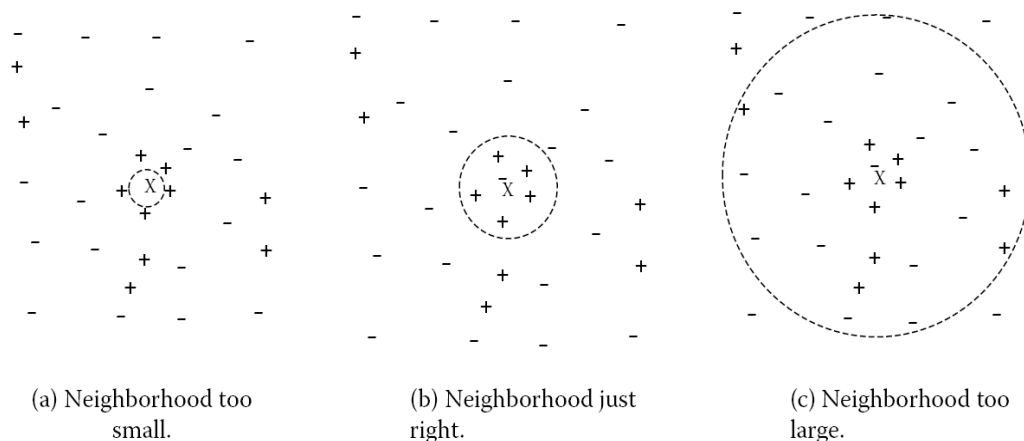


FIGURE 6.1: Choosing k value in kNN.

Basic kNN Algorithm

Input : the set of objects taken for training 'D' or reference. Test object is 'Z' which is a vector of attribute values and set of classes used to label the objects L

Output: is reflected by $Cz \in L$, the class of 'z' .

for each object $Y \in D$

do

|Compute $d(z, y)$, the distance between z and y;

end

7. ARTIFICIAL NEURAL NETWORK

Neural Networks are large networks of simple processing elements or nodes which process information dynamically in response to external inputs [14]. The nodes are simplified models of neurons. Artificial Neural Network (ANN) or Neural Network is a mathematical model inspired by biological neural networks. This network of inter connected artificial neurons process information using a connectionist approach to computation. During the learning phase, in most cases, a neural network is an adaptive system that changes structure. Neural networks are used to model between inputs and outputs or to find patterns in data.

The interconnections between the neurons in the different layers of each system is referred to by the word network in the term 'artificial neural network'. The input neurons of the first layer send data via synapses to the third layer of output neurons. More complex system have more number of layers. The synapses store parameters called "weights" that operate on data in calculations.

An ANN is typically defined by 3 parameters:

1. The inter connection pattern between different layer of neurons
2. The learning process for updating the weights of inter connection
3. The activation function that converts neuron weighted input to output activation.

ANN's ability to be used as an arbitrary function approximation mechanism that 'learns' from observed data is probably the greatest advantage of ANN. It is however essential to use them for more complex and relatively good understanding of underlying theory.

1. Model Choice: Depends on data representation and application. A very complex model tends to lead to problem with regard to learning.
2. Learning algorithm: Many trade-offs exists between learning algorithms. With correct parameters for training on a fixed dataset almost any algorithm will work well. However significant amount of experimentation is required for selecting and tuning an algorithm for training on unseen data.
3. Robustness: ANNs can be made robust by choosing the correct model, appropriate cost function and proper learning algorithm.

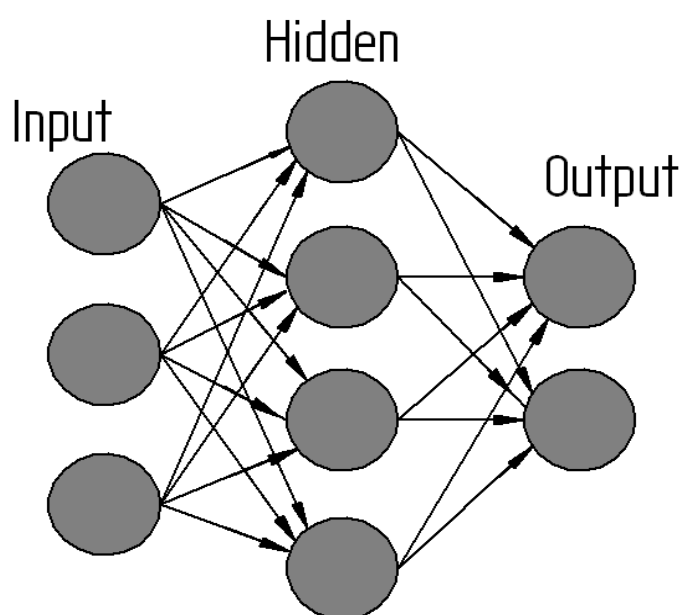


FIGURE 7.1: Architecture of Artificial Neural Network (ANN).

7.2 Training Using Back Propagation Algorithm

Back propagation learning emerged as the most significant result in the field of artificial neural network (15). Back Propagation Network (BPN) is a multilayered, fully connected feed forward network and uses supervised mode of learning. After an input pattern has been applied to the first layer of network units, it is propagated to each layer and an output is generated. This output is compared with desired output (target) and an error signal computed for each output unit. The error signal is transmitted backward from the output layer to each node in the hidden layer. This process repeats, layer by layer until each node in the network has received an error signal the describe its relative contribution to the total error. Based on the error signal received, connection weights are then updated by each unit to cause the network to converge towards a stage that allows all the training patterns to be encoded.

The back propagation network chosen in the work has two hidden layer one output layer and one input layer. The network parameters chosen are:

- Learning Rate = 0.05
- Number of Hidden Layer = 2

- Number of Neuron in the 1st Hidden Layer = 10
- Number of Neuron in the 2nd Hidden Layer = 10
- Number of Neuron in the Input Layer = 2
- Number of Neuron in the Output Layer = 3

We have used following function \s for training ANN

- Transfer Function: Log Sigmoid

It takes around 1500 iterations to convergence. However the number of iterations change whenever a new ANN with new architecture is chosen.

7.3 Implementation Using BPA

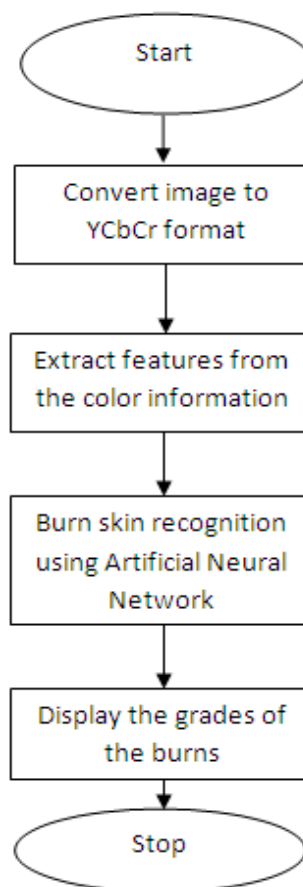


FIGURE 7.3: Flow Chart for Training ANN With Burn Images Using BPA .

8. RESULTS

8.1 Template Matching Method

This is the simple method of classification, a new test vector containing mean and DCT values of sample image is matched with the previously stored templates belonging to different grades. Depending on the closeness of the match, test sample is categorized. Testing of 120 images, with 40 images from each category is done and tabulated in Table 8.1. Twenty images from each category are used for training and remaining 20 for testing with cross validation.

Types of Image	No. of Images	Correctly Classified	Misclassified as			Efficiency as in %
			Grade 1	Grade 2	Grade 3	
Grade 1	40	22		6	12	55
Grade 2	40	29	4		7	72.5
Grade 3	40	28	9	3		70

TABLE 8.1: Template Matching Method.

8.2 kNN Classifier

In this method, classification is based on nearest neighbor method. Depending on the value of the k specified, algorithm checks for the k nearest neighbors of training values, surrounding a test value. The closest neighbor determines the class of test value. In this work, multi class kNN classifier is used with k value as 3. Cross validation of training and testing vectors are done by using 3 fold technique by randomly selecting training and testing vectors, the results of the classifier is shown in Table 8.2

Types of Image	No. of Images	Correctly Classified	Misclassified as			Efficiency as in %
			Grade 1	Grade 2	Grade 3	
Grade 1	40	27		4	9	67.5
Grade 2	40	33	5		2	82.5
Grade 3	40	30	7	3		75

TABLE 8.2: kNN Classifier.

8.3 ANN Classifier

In this method of classification Back Propagation Algorithm is used to train the ANN. Best efficiency is obtained when the 1st hidden layer contains 10 neurons and 2nd hidden layer contains 10 neurons. It takes 800 iterations to converge, however number of iterations change whenever the network is trained with different number of layers and neurons. ANN Classifier has resulted with an efficiency of more than 90% as shown in table 8.3:

Type of Image	No. of Images	Classified as			Efficiency in %
		Grade1	Grade2	Grade3	
Grade1	40	38	2	0	95
Grade2	40	0	39	1	97.5
Grade3	40	0	1	38	95

TABLE 8.3: ANN Classifier.

9. CONCLUSION

In this paper classification of the burn is achieved based on the depth of burn which is extracted from the color characteristics of burn images. The performance efficiency of three classifiers have been compared. It is seen that the ANN method of analyzing the Skin Burn Image reflects the grade of the injury very close to that by a Clinician. The features considered for classification are Mean and DCT. These features can be computed easily without any complex algorithm and thus reduces classification time. Burn Images are fed into the system and the grade of the wound would be displayed on the screen after the classification algorithm processes the image. Algorithms proposed

are simple and thus result can be computed very fast. Using Internet and mobile phone the work can be converted to a tele medicine project.

ACKNOWLEDGMENT

Authors are thankful to Sri Devraj Urs Medical Center for providing the burn wound photographs . This help, cooperation and services are highly appreciated and acknowledged. Our thanks also due to Dr. Jyothi Subash, Dermatologist, Bangalore and for her views on this topic and we wish to express sincere thanks to Mr. Deepak L Student of Atria Institute of Technology for his support on this project.

10. REFERENCES

1. Clarke, J.A.: A Color Atlas of Burn Injuries. Chapman & Hall medical, London, (1992).
2. Serrano, C., Rao, L.M., Acha, B.: Evaluation of a Telemedicine Platform in a burn Unit. Proc. IEE Int. Conf. on Information Technology Applications in Biomedicine, Washington DC (USA) (1998) 121-126.
3. Roa, L.M., Gomez-Ciz, T., Acha, B., Serrano, C.: Digital Imaging in Remote Diagnosis of Burns. Burns, 7 (1999) 617-624.
4. Begona Acha, Carmen Serrano, Sergio Palencia, Juan Jose Murillo. Classification of burn wounds using support vector machines, (2001).
5. R. Gonzalez and R. Woods, "Digital image processing," IEEE Transactions on Systems, Man, and Cybernetics, 2008.
6. Mathworks. (2010) Image Processing User guide. [Online]. Available: <http://www.mathworks.com/help/pdf>
7. Afromowitz, M.A., Van Liew, G.S., Heimbach, D.M.: Clinical Evaluation of Burn Injuries Using an Optical Reflectance Technique. IEEE Trans. on Biomedical Engineering, 2 (1987) 114-127.
8. Afromowitz, M.A., Callis, J.B., Heimbach, D.M., DeSoto, L.A., Norton, M.K.: Multispectral Imaging of Burn Wounds: A New Clinical Instrument for Evaluating Burn Depth. IEEE Trans. on Biomedical Engineering, 10 (1988) 842-850.
9. Hansen, G.L., Sparrow, E.M., Kokate, J.Y., Leland, K.J., Iazzo, P.A.: Wound Status Evaluation using Color Image Processing.
10. Serrano, C., Acha, B., Acha, J.I.: Segmentation of Burn Images based on Color and Texture Information. SPIE Int. Symposium on Medical Imaging, 5032, (2003) 1543-1550, San Diego (CA, USA).
11. Image classification based on color and texture analysis. First intel workshop on Image and Signal Processing and Analysis June 14-15, 2000, Pula, Croatia.
12. J E Bennett R O Kingman "Evaluation of Burn Depth by the use of Radioactive Isotope- by An experimental study ", Ploic and reconstructive surgery, vol. 20, No.4, PP 261-272, 1957.
13. R. Duda, P. Hart, and D. Stork, "Pattern classification, ed, 2001.
14. Khalil Shihab, "A Back propagation Neural Network for Computer Network Security" in proceedings of Journal on computer Science, vol. 2, no.9, 2006, pp: 710-715.
15. Artificial Neural Network R.Yagna Narayana 2005.
16. Thira, J.P., Macq.B.: Morphological Feature Extraction for the classification of Digital Images of cancrus Tissues. IEEE.

Target Detection by Fuzzy Gustafson-Kessel Algorithm

Mousumi Gupta

*Assistant Professor/Comp Sc & Eng Department
Sikkim Manipal Institute of Technology
Gangtok, 737136, India*

mousmi_gt@yahoo.co.in

Abstract

Many commercially available radar systems offer a range of filter options but the problem of clutter rejection for target detection is still present in a number of situations. Rejection of clutter and detection of targets from radar captured data is a challenging task. Raw data captured by radar are not always scaled. A normalization technique has been proposed which transforms the radar captured data into 8 bit. As 8 bit data is easy to analyze and visualize. A modification on Fuzzy c-means has been done by developing Fuzzy Gustafson–Kessel (FGK) algorithm and the result shows robustness of this proposed method.

Keywords: Target Detection, Clutter Rejection, Data Normalization, Fuzzy Clustering, Fuzzy Gustafson-Kessel (FGK).

1. INTRODUCTION

Detection of targets from radar data is a desired task in automatic target recognition (ATR). Interaction of electro-magnetic scattering between the target of interest and the rough surface makes target detection more challenging. Clutter rejection and target detection is of interest in recent years because of the need for object detection on the sea surface, remote sensing of vegetation for crop production assessment, and many other applications. There are several existing method has already been proposed for radar target detection [1-4].

In ISAR (Inverse Synthetic aperture Radar), the target rotates and the radar is stationary. Target images can be obtained by transmitting wideband signals, and high cross range resolution is obtained by coherently accumulating number of echoes from different aspect angles [5]. The goal of radar imaging system is to detect the targets particularly for surveillance. Clutter is the major problem for radar operations. Clutter refers to radio frequency (RF) echoes returned from targets which are uninteresting to the radar operators. The nature of clutter varies with application and radar parameters [6] because of many users and the over-crowding of the spectrum; electromagnetic interference is a common occurrence with current communication and electronic equipments. Hence, analysis performed to either avoid or eliminate such interference, which is termed as clutter. There are several works has been carried out with fuzzy c-means clustering [7] [8] but none of them has been able to reject clutter from the data. Lejiang et al[9] has proposed a fuzzy c-means clustering for duplicate data cleaning. Fuzzy Gustafson-kessel has been used by Niladri et al[10] for remote sensing change detection. Our method fuzzy Gustafson-kessel not only detects targets but reject clutter.

The main task of a radar signal processor is to make decisions whether a target present or not. After a signal has been transmitted, the receiver starts receiving return signals, with those originating from near objects arriving first because time of arrival translates into target range. The signal processor places a raster of range bins over the whole period of time, and now it has to make a decision for each of the range bins as to whether it contains an object or not. This decision-making is severely hampered by noise. Atmospheric noise enters into the system through the antenna, and all the electronics in the radar signal path produces noise too. In this

paper a data scaling method has been proposed. Radar captured data even if used as a form of images but these images do not normally resemble to those produced by conventional imaging system. The large volume of image data would overwhelm the available image analysis capabilities so scaling of captured data is becoming popular for target detection and other image analysis task [11] [12]. The raw data captured by the radar are all complex in nature and are not scaled. For an unscaled matrix, retrieving information is a very difficult task. In this proposed approach first we have used a method to normalize the captured data by dividing each data with its mean value. The matrix is scaled in such a way that the pixel ranges from 0 to 255, as this is the range generally used for gray level image processing. Scaling of the radar data is able to reject small portion of clutters. Fuzzy clustering has been done on the scaled matrix by considering that the targets and the clutters are distinguishable significantly.

Fuzzy c-means clustering generally employs Euclidean norm to measure the dissimilarity between patterns and cluster centers. Only spherical clusters can be detected properly using it. Fuzzy Gustafson and Kessel introduced [13] adaptive distance norm to measure the distance between clusters using fuzzy covariance matrix (a fuzzy equivalent of the classical covariance) - a representation of cluster centers along with data points. Using FGK [13] clusters with any shape can be detected. In case of radar captured data targets are not always in spherical structure. And so by normal fuzzy c-means clustering detection of targets are not always possible. To overcome from this drawback in this paper we proposed fuzzy Gustafson-Kessel for target detection on scaled data.

2. DATA CAPTURING

A MATLAB simulated environment has been created where four points scatterer has been designed in the model. Where the left and right point scatterer is rotating with an angular step of 0.3° about a central axis the fourth one is the rear point scatterer. The rear side point scatterer is placed 17m behind the middle point scatterer and at a 0.9m cross range distance with respect to the right side point scatterer. The cross range distance between left and middle point is 0.5m. Similarly the middle and right points are separated in cross range by 0.5m. By this model the captured matrix is of size 21×724 . As the point scatterer assembly is rotated starting from 23.5° to 29.5° with a step of 0.3° . In this way we have generated 21 row vectors each with size 1×4 . So, for a single Radio Frequency, a data matrix will be generated having the dimension of 21×4 . The numbers of relative frequencies (i.e., RF) are taken as 181 because the RF sweep range is 1.7GHz to 2.6GHz with a step size of 5MHz. Finally 181 number of 21×4 data matrix has been generated through the simulation model and the resulting data matrix on which the work has been carried out is with dimension $[21 \times (181 \times 4)]$ i.e., 21×724 .

3. SCALING OF DATA

Constructing an ISAR image requires data collection in both frequency and angular dimensions. ISAR imaging systems produce electromagnetic images of targets in the range-Doppler domain. If the data are evenly sampled and the sampling rate is dense enough and if the total angular looks on the target is small then an ISAR can be obtained by using a 2 dimensional FFT algorithm. The data captured through ISAR are always having long float values. In this paper we have converted the captured value first into short floating point value by dividing each pixel with its mean value. And the scaling operation has been done by using the following criteria;

```

If max (pixel_value)>255
then new_pixel=max(pixel_value)-255
else
new_pixel=255-max
end

```

4. FUZZY GUSTAFSON-KESSEL ALGORITHM (FGK)

Fuzzy cluster analysis allows gradual memberships of data points to clusters measured as degrees in [0,1]. This gives the flexibility to express that data points can belong to more than one cluster. These membership degrees offer a much finer degree of detail of the data model. Aside from assigning a data point to clusters in shares, membership degrees can also express how a data point should belong to a cluster. The concept of these membership degrees is substantiated by the definition and interpretation of fuzzy sets. A fuzzy cluster model of a given data-set X into c clusters is defined to be optimal when it minimizes the objective function;

$$J_f(X, U_f, C) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2 \quad (1)$$

Where $X = \{x_1, x_2, \dots, x_n\}$ be the set of pixels and c is the number of clusters ($1 < c < n$) represented by fuzzy sets μ_{Γ_i} and $i = 1, 2, \dots, c$. U_f is the partition of X if it satisfy the following;

$$\sum_{j=1}^n u_{ij} > 0 \quad \forall i \in 1, 2, \dots, c \text{ and}$$

$$\sum_{i=1}^c u_{ij} = 1 \quad \forall j \in 1, 2, \dots, n$$

'd' is the Euclidean distance between cluster centers and the data points. But this distance only makes it possible to identify spherical clusters.

The Gustafson–Kessel algorithm [13] replaces the Euclidean distance by a cluster-specific Mahalanobis distance, so as to adapt to various sizes and forms of the clusters. The Gustafson–Kessel algorithm tries to extract much more information from the data than the algorithms based on the Euclidean distance. In the Gustafson–Kessel algorithms the clusters are parameterized by the mean (center) and the distance norm matrix (related to the covariance matrix of the cluster). The Gustafson–Kessel algorithm tries to extract much more information from the data than the algorithms based on the Euclidean distance. It is more sensitive to initialization, therefore it is recommended to initialize it using a few iterations of FCM depending on the considered partition type. The Fuzzy Gustafson–Kessel algorithm exhibits higher computational demands due to the matrix inversions.

For a cluster i, its associated Mahalanobis distance is defined as;

$$d^2(x_j, C_i) = (x_j - c_i)^T \sum_i^{-1} (x_j - c_i)$$

Where \sum_i is the covariance matrix of the cluster.

5. RESULTS

The raw data captured by the radar is complex values. Where the minimum value is 9.8868e-08 and the maximum value is -6.8941e-04. Data has been scaled by using the proposed scaling method which is described in algorithm 1.

- Step 1: data= R ;
- Step 2: m = mean(R);
- Step 3: MAT1= abs(int($\frac{R}{m}$))

```

Step 4: MAXIMUM= MAX (MAT1);
Step 5: if MAXIMUM ≥ 255 then
Step 6: VARIABLE1=MAXIMUM-255;
Step 7: if MAXIMUM ≤ 255 then
Step 8: VARIABLE1=255-MAXIMUM;
Step 9: FINAL_MATRIX= MAT1-VARIABLE1;
Step 10: if FINAL_MATRIX ≤ 0 then
Step 11: FINAL_MATRIX=0;
    
```

Algorithm 1: Proposed data scaling method

Figure 1(a) contains contour plot for the original complex data matrix. Figure 1 (b) is the generated contour plot after applying proposed algorithm. By which we can easily determine that there are three targets present in the environment.

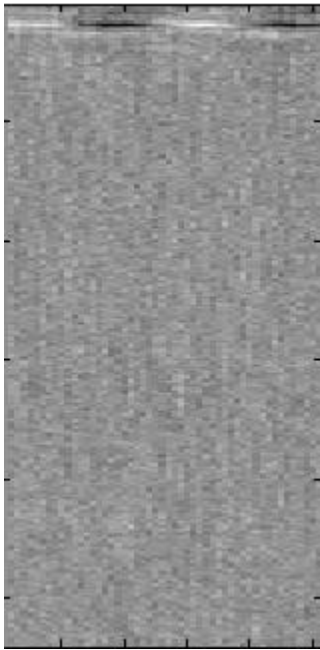


FIGURE 1(a)

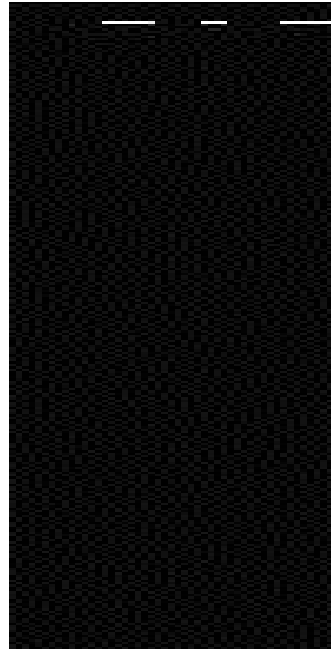


FIGURE 1(b)

All three targets are taken as flat plates for experiment with approximate dimensions (.56mX.56m) for left target, (.3mX.3m) for middle (.71mX.71m) for right target. The Gustafson–Kessel algorithm models each cluster Γ_i by both its center c_i and its covariance matrix \sum_i where $i = 1, 2, \dots, c$. Thus cluster prototypes are tuples $C_i = (c_i, \sum_i)$ and both c_i and \sum_i are to be learned. The eigen structure of the positive definite $p \times p$ matrix \sum_i represents the shape of cluster i .

The cluster center c_i has been calculated by using the formula;

$$c_i = \frac{\sum_{j=1}^n u_{ij}^m x_j}{\sum_{j=1}^n u_{ij}^m}$$

The membership degree has been calculated by;

$$u_{ij} = \frac{d_{ij}^{-\frac{2}{m-1}}}{\sum_{l=1}^c d_{lj}^{-\frac{2}{m-1}}}$$

The covariance matrix has been calculated by;

$$\sum_i = \frac{\sum_i^*}{\sqrt[p]{\det(\sum_i^*)}}$$

$$\text{Where } \sum_i^* = \frac{\sum_{j=1}^n u_{ij}(x_j - c_i)(x_j - c_i)^T}{\sum_{j=1}^n u_{ij}}$$

6. CONCLUSION

This work describes a natural extension of Fuzzy c-means clustering with Gustafson-Kessel distance method. The new formulation should prove extremely beneficial to improving the performance of fuzzy clustering algorithms in the field of target detection and clutter rejection. Since the pixels differences between the clutter and target is very low so they are not separable by sharp boundaries (as they are highly overlapped), fuzzy Gustafson-Kessel algorithm seem to be more appropriate and realistic choice to separate them.

7. REFERENCES

1. D. Colak, R. J. Burkholder, and E. H. Newman, Multiple sweep method of moments analysis of electromagnetic scattering from 3D objects on ocean-like rough surfaces, *Microwave and Optical Technology Letters*, vol. 49, pp. 241-247, 2007.
2. L. Guo and K. Cheyong, Light scattering models for a spherical particle above a slightly dielectric rough surface, *Microwave and Optical Technology Letters*, vol. 33, pp. 142-146, April 2002.
3. Steven P Jacobs, Joseph A. O'Sullivan, Automatic target recognition using sequences of high resolution radar range profiles, *IEEE transactions on aerospace and electronic systems*, Vol 36, No 2, 2000.
4. Rajesh K, Radar target detection in Weibull clutter by adaptive filtering with embedded CFAR, *IEEE Electronics Letters*, Vol 35 , 597-599, 1999.
5. M. Gupta et al. *Pattern Recognition Letters*, Vol 33, pp. 1682-1688, 2012.
6. Ilteris Demirkiran, Donald D. Weiner and Andrew Drozd , Effect of In-band Intermodulation Interference on Direct-Sequence Spread Spectrum (DSSS) Communication Systems for Electromagnetically Diverse Applications, 2007 IEEE.
7. Gang Wang , Jinxing Hao , Jian Ma, Lihua Huang , A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering, *Expert Systems with Applications*, 2010.
8. Dzung L. Pham, Spatial Models for Fuzzy Clustering, *Computer Vision and Image Understanding*, Vol 84, pp. 285-297 , 2001.

9. Tara.Saikumar, B.K.Anoop, P.S.Murthy, "Tara Kernel Fuzzy Clustering (TKFCM) for a Robust Adaptive Threshold Algorithm based on Level Set Method", International Journal of Information Technology Convergence and Services, Vol.2, No.1, February 2012.
- 10.Ashish Ghosh , Niladri Shekhar Mishra , Susmita Ghosh, Fuzzy clustering algorithms for unsupervised change detection in remote sensing images, Information Sciences,2010.
- 11.Guangzhi Cao et al , The Sparse Matrix Transform for Covariance Estimation and Analysis of High Dimensional Signals, IEEE Trans. on Image Processing, Vol 20, pp.625-640, 2011.
- 12.Leonardo R et al.,Fast Signal Analysis and Decomposition on Graphs using the Sparse Matrix Transform, in the Proceedings of the International Conference on Acoustic, Speech, and Signal Processing , March 14-19, 2010.
- 13.Gustafson, E. E. and Kessel , W. C., Fuzzy clustering with a fuzzy covariance matrix Proc. of the IEEE Conference on Decision and Control, San Diego, pp. 761–766. IEEE Press, Piscataway, NJ.1979.

HABIT: Handwritten Analysis Based Individualistic Traits Prediction

Abdul Rahiman M

*Director, AICTE
Ministry of HRD, Govt of India
New Delhi, INDIA*

rehman_paika@yahoo.com

Diana Varghese

*Business Analyst, Mu Sigma Business Soln Pvt Ltd
Bangalore, INDIA*

dianavarghese100@gmail.com

Manoj Kumar G

*Associate Professor, Dept of Computer Science
LBSITW, Trivandrum, Kerala, INDIA*

manojkumar_gg@hotmail.com

Abstract

Handwriting Analysis is a scientific method of identifying, evaluating and understanding an individual's personality based on handwriting. Each personality trait of a person is represented by a neurological brain pattern. Each of these neurological brain patterns produces a unique neuromuscular movement that is the same for every person who has that particular personality trait. When writing, these tiny movements occur unconsciously. Strokes, patterns and pressure applied while writing can reveal specific personality traits. The true personality including emotional outlay, fears, honesty and defenses are revealed. Professional handwriting examiners called graphologists analyze handwriting samples for this purpose. However, accuracy of the analysis depends on how skilled the analyst is. The analyst is also prone to fatigue. High cost incurred is yet another deterrent. This paper aims at implementing an off-line, writer-independent handwriting analysis system "HABIT" (Handwriting Analysis Based Individualistic Traits Prediction) which acts as a tool to predict the personality traits of a writer automatically from features extracted from a scanned image of the writer's handwriting sample given as input. The features include slant of baseline, pen pressure, slant of letters and size of writing. The implementation uses Java and Eclipse-Indigo as tools.

Keywords: Handwriting Analysis, Feature Extraction, Patterns.

1. INTRODUCTION

Handwriting Analysis, also known as Graphology is a scientific method of identifying, evaluating and understanding personality through the strokes and patterns revealed by handwriting. Handwriting reveals true personality including emotional outlay, fears, honesty, defenses and many other individual traits. Handwriting is often referred to as brain writing. Each personality trait is represented by a neurological brain pattern. Each neurological brain pattern produces a unique neuromuscular movement that is the same for every person who has that particular personality trait. When writing, these tiny movements occur unconsciously. Each written movement or stroke reveals a specific personality trait. Graphology is the science of identifying these strokes as they appear in handwriting and describing the corresponding personality trait.

Personality identification of a human being by their handwriting is an old technique. Earlier this was done manually by spending a lot of time to predict the nature of the person. Handwriting is brain writing, representing the mental status of the person. Handwriting analysis is a projection technique that profiles the human behavior in areas of social skills, achievements, thinking styles

and work habits. Handwriting also depicts the possible ways of a person's transactions with stress.

In case of manual analysis, accuracy of the analysis depends on how skilled the analyst is. The analyst is also prone to fatigue when several samples are to be analyzed. High cost incurred in getting the aid of a well-experienced graphologist is yet another deterrent.

Compared to manual analysis, automated handwriting analysis is very fast and accurate in the prediction of human personality. Collecting digital samples of handwriting and using computer prediction is a very low-cost and convenient method. One can easily give the digital sample of his/her handwriting to a computer and it calculates the features using image processing techniques and predicts the nature of the writer.

In this paper, a system is proposed to predict the personal behavior of an individual from their handwriting analysis in digital form. To predict the actual personality of the individual there are various features, such as slant, size, pressure, upper zone (as in **l, t, h, S**), lower zone (as in **g, q, y, z**), word spacing, line spacing, page margins, middle zone or case (as in **a, o, c, s, e**). In the proposed system the features utilized are the slant of letters, slant of the baseline, pen pressure and size of the letters as they are enough to predict the behavior of the person.

2. DEVELOPMENT OF GRAPHOLOGY

Graphology is the study of a person's personality from their handwriting. It has a long and respectable history dating as far back as the ancient Romans. The historian Seutonius Tranquillus was among the first to notice in the handwritings of several emperors he was researching that the letters were formed differently in each case. He theorized that this could indicate different character traits. The 19th century work by Abbott Jean-Hippolyte Michon and Abbott Flandrin, in which they collected innumerable handwriting samples and matched styles of writing with various temperaments, was the beginning of modern graphology. It was Abbott Michon who coined the term 'graphology'.

The next great leap in the scientific evaluation of handwriting came from Ludwig Klages who was the first to create a complete and systematic theory of 'graphology'. Klages classified personality traits by evaluating the up and down strokes of handwriting. For example, certain handwriting rhythms would indicate someone's intellectual passion, whereas certain letter forms would display someone's sense of inferiority. A 'normal' person would have a balance of contraction and release, whereas unstable people would have an unbalanced rhythm. Even though the system was developed, there did not exist an objective way of rating someone's entire personality.

Joseph Zubin and Thea Lewinson built upon Klages work and developed a system of scales, called the L-Z scales that evaluated the quantitative and qualitative aspects of handwriting. According to Lewinson-Zubin, there are four special components of handwriting – vertical, horizontal, depth, and form – by which each written letter can be evaluated. These four components yield the following dimensions of personality: rational, social, emotional and instinctual.

The vertical component concerns the height of the middle zone of a letter which emphasizes self-importance. The direction of the vertical lines belies the individual's mood level. The horizontal component is measured by the distance between letters and words. Right/left slants are also included in this measurement. Horizontal traits measure the relationship between the individual and his or her environment. The depth component is the pressure of the writing which represents one's instinctual drives. The form component measures the contour of the writing which can signify the degree of one's creativeness.

3. SYSTEM DESCRIPTION

Professional handwriting examiners called graphologists often identify the writer with a piece of handwriting. Accuracy of handwriting analysis depends on how skilled the analyst is. Although human intervention in handwriting analysis has been effective, it is costly and prone to fatigue.

The proposed system HABIT focuses on developing a tool for behavioral analysis which can predict the personality traits automatically with the aid of a computer. The system is designed to analyze scanned images of handwritten documents. The images are converted into binary black and white pixel images. It is difficult for the program to identify different letters, words and lines by pixel analysis without human intervention. Therefore, human decision process guides the system to achieve higher accuracy. Here the aim is to create software that allows a user to analyze handwriting samples, with the aim of making the process faster and more objective.

HABIT is an off-line, writer independent handwriting analysis system that predicts the personality of a person. The input to the system is a scanned image of a handwriting sample of the writer. The behavioral analysis is done from the baseline slant, the pen pressure, the slant of the writing and size of letters. The output is a set of personality trait of the writer. The entire system is depicted in figure 1.

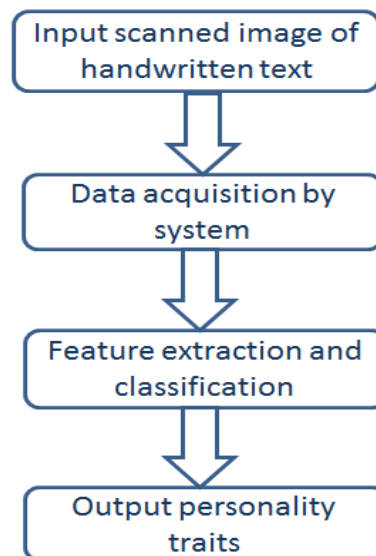


FIGURE 1: HABIT System- An overview.

4. IMAGE ACQUISITION & PREPROCESSING

Images of handwritten samples are uploaded to the system. The system requires preprocessing work from the user in order to begin calculating the scales from horizontal and vertical category. The uploaded image is preprocessed and resized to the correct orientation. Steps in Image acquisition and feature extraction are shown in figure 2.

The application allows users to crop images into lines, words and characters. Once the images are cropped, the cropped images will be displayed on the scratchpad. In the analysis window, a cropped image will be loaded.

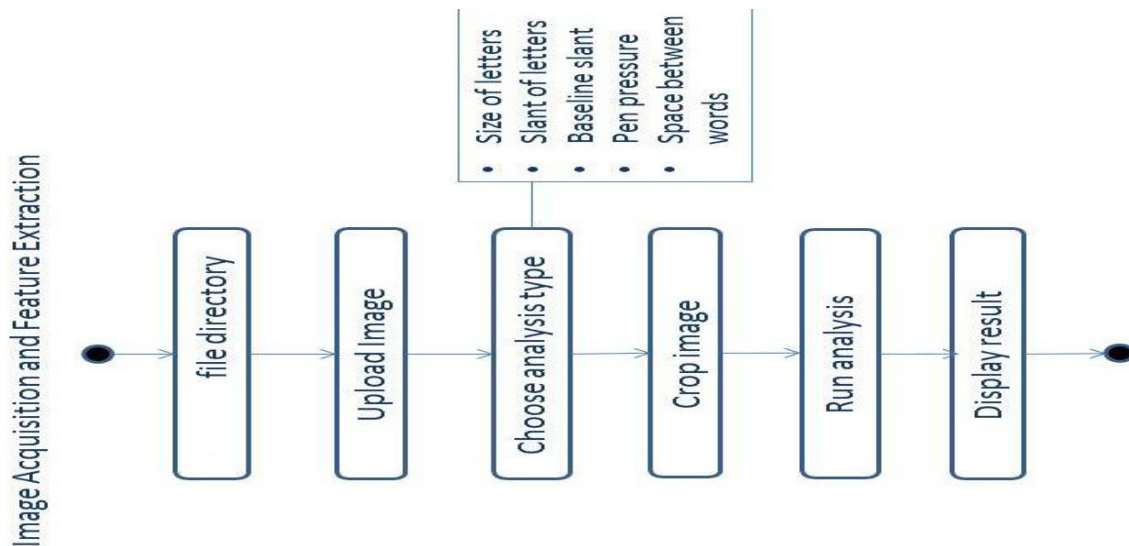


FIGURE 2: Image Acquisition & Processing.

Depending on the specimen chosen data points can be plotted. The user must plot some data points on the image that identify the top, middle, base and bottom lines of the handwriting. The collection of the data points themselves has a subjective nature because they depend on the user's expert eye to select points of interest. Due to this reason, linear regression approximation of least squares is a suitable method.

5. FEATURE EXTRACTION STAGE

5.1 Pen Pressure

One of the most important features in a handwriting sample is the pressure of writing. The amount of pressure exerted on the paper while writing indicates the depth of feeling, also called emotional intensity, of the writer. Based on the pen pressure, the writer can be classified as a light writer, medium writer or heavy writer. For the analysis of pen pressure, first the scanned image is converted into a grey scale image. Mean grey level value is computed using the grey level values of the image pixels. This mean grey level value of the image is compared with the pre-determined threshold value, th_0 . Higher value of the mean indicates lighter pressure. If the mean is less than th_0 , then the writing pressure is considered to be high. Such a person has very deep and enduring feelings. This writer may forgive, but he will never forget. He feels situations intensely. If mean is greater than th_0 , then writing pressure is considered to be light. Such a person can endure traumatic experiences without being seriously affected. Emotional experiences do not make a lasting impression on him.

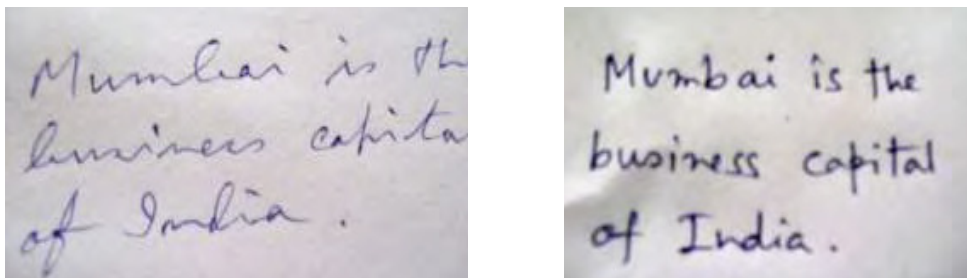


FIGURE 3: Light writer & Heavy writer

5.2 Slant of letters and Slant of baseline

The baseline in one's handwriting reveals a lot of accurate information about the writer. Baseline in one's handwriting is the line along which the writing flows. The three most common baselines found in any handwriting are ascending, descending and level. Given a set of data points, we need a trend line that passes closely among the points, especially if there is large number of data points. This trend line that we compute is called regression line. This regression line is computed using Least-Squares Linear Regression. To calculate the slant in drawn lines, the formula is

$$\theta = \tan^{-1} \frac{(y_2 - y_1)}{(x_2 - x_1)}$$

Standard reference angle (θ_0) is considered to be 90. θ is compared with θ_0 to classify the slant height.

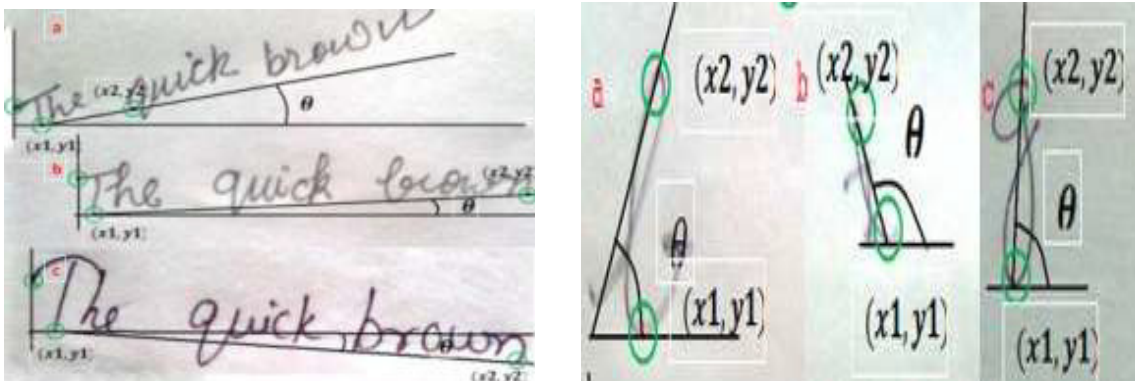
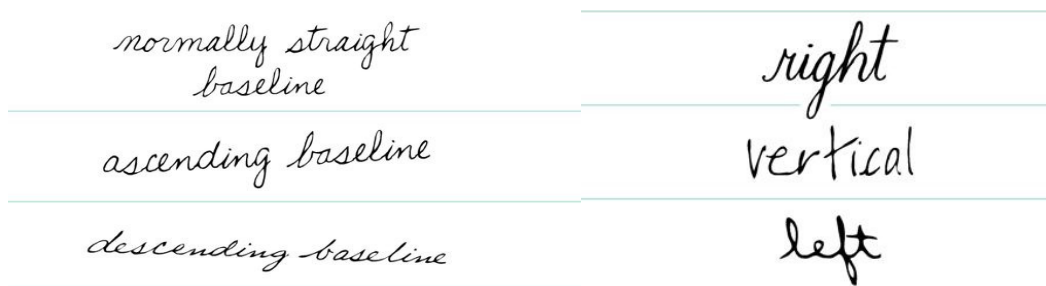


FIGURE 4: Slant of Letters & Trend line.

5.3 Size of Letters

The size of handwriting is judged by a benchmark of 3mm as normal writing and full height of 9mm. Other than this, writing is classified as large or small writing. The letters are divided into 3 zones: lower, upper and middle. Large writing can portray someone who is superior, takes pride, outgoing and extrovert, arrogant, boastful or they put on an act of confidence. Small handwriting can mean people who are respectful, tolerable, introvert, shy, deep thinkers or academic. If letters which reach into the upper zones are very extended then the person probably has unrealistic expectations of what they can truly achieve. Wide upper zone loops can show people who are deep thinkers and dreamers; then those whose letters go up then down directly over themselves are often un-imaginative. If the letters in the lower zone are straight it shows people who like to get the job done; those who loop them are often full of energy, good at making and investing money, or/and need security. The different zones are shown in figure 5. To find the size of letters we use the Pythagora's theorem to find the distance between the top point and baseline. For this we use the formula:

$$\sqrt{(\Delta x^2 + \Delta y^2)} = c^2 \quad \text{where } \Delta x \text{ is the distance between x co-ordinates}$$

Δy is the distance between y co-ordinates

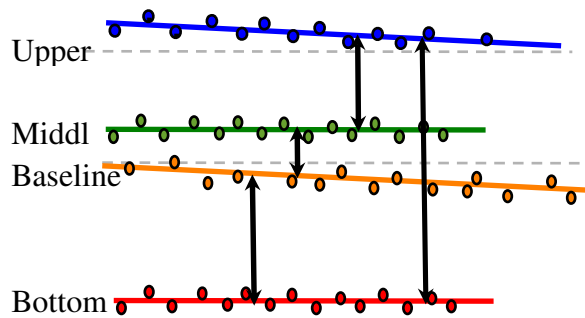


FIGURE 5: Different Zones.

5.4 Spacing Between Words

A person who has some words widely spaced are often open, honest but deep in thought and people with words narrow spaced may be unstable in either emotions or thinking. Spacing between the words is shown in figure 6. To find the distance between the two words perpendiculars are dropped from the end point of the first word and starting point of the second word. Length of the hypotenuse joining the two perpendiculars gives the distance between the two words. The equation to compute the length of the hypotenuse is

$$\sqrt{(\Delta x^2 + \Delta y^2)} = c^2$$

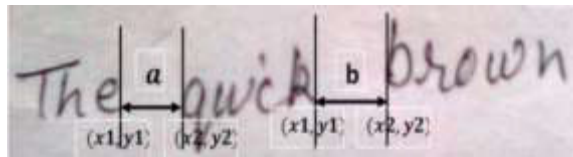


FIGURE 6: Word Spacing.

5.5 Handwriting Style Classification

The handwriting style classification is shown in Table 1 below.

Sl. No.	Writing Category	Sub-category Criteria
1	Size of letters	Large: size > normal size Small: size < normal size Medium : size = normal size
2	Slant of letters	Right: $\theta > \theta_0$ Left: $\theta < \theta_0$ Vertical: $\theta = \theta_0$
3	Baseline	Raising: $\theta > 0$ Falling: $\theta < 0$ Straight: $\theta = 0$ Erratic: otherwise
4	Pen Pressure	Light Pen: threshold > th0

		Heavy Pen: threshold < th0
5	Space between words	Far: space > average space Near: space < average space

TABLE 1: Handwriting Styles & Classification.

5.6 Personality Traits Predicted by Various Handwriting Styles

Writing Categories	Psychological Personality Behaviour
Large Letters	Likes being noticed, stands out in a crowd
Small Letters	Introspective, not seeking attention, modest
Medium Letters	Adaptable, fits into a crowd, practical, balanced
Right Slant	Sociable, responsive, interested in others, friendly
Left Slant	Reserved, observant, self-reliant, non-intrusive
Vertical	Practical, independent, controlled, self-sufficient
Light Pen Pressure	Can endure traumatic experiences without being seriously affected. Emotional experiences do not make a lasting impression
Heavy Pen Pressure	Have very deep and enduring feelings and feels situations intensely.
Raising Baseline	Optimistic, upbeat, positive attitude, ambitious and hopeful
Falling Baseline	Tired, overwhelmed, pessimistic, not hopeful
Straight Baseline	Determined, stays on track, self-motivated, controls emotions, reliable, steady
Erratic Baseline	Wavering, lacks definite direction, emotionally unsettled, unpredictable
Far Spaced Words	Desires more space, enjoys privacy
Close Spaced Words	Closeness of sentiment and intelligence

TABLE 2: Personality Traits.

6. SYSTEM IMPLEMENTATION

The input to the system is a scanned image of a handwriting sample of the writer. The behavioral analysis is done from the baseline slant, the pen pressure, the slant of the writing size of letters and spacing between words. The output is a set of personality traits of the writer. The method used to implement this is simple linear regression which is an approach to modeling the relationship between a scalar dependent variable y and an explanatory variable denoted x .

6.1 Least Squares Linear Regression

Predictions for y from each value of x in the data will usually differ from the actual value of y that is being predicted. If you square the difference and add up these squared differences across all the predictions, you get a number called the residual or error sum or squares. With any two variables x and y , there exists one formula that will produce the best, or most accurate predictions for y given x . Any other equation would not fit as well and would predict y with more error. That equation is called the least squares regression equation. Step in running linear regression are:

1. Specimen is displayed on the scratchpad
2. User plots data points on the sample
3. Regression line computed using selected data points

Formula for line-approximation using n points,

$$a_1 = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

$$a_0 = \bar{y} - a_1 \bar{x}$$

where a_1 is the line slope and a_0 is the line intercept .
The regression line obtained is plotted in figure 7.

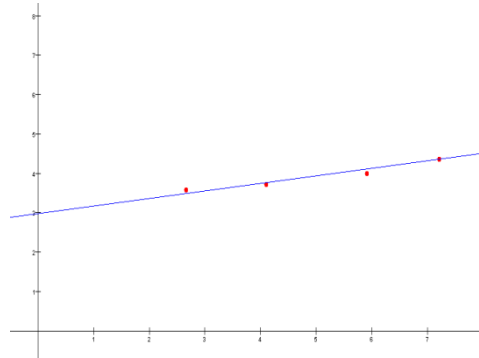


FIGURE 7: Regression line of the plotted points

Hibernate has been used to perform operations on the database like select, insert, update and delete the records in the table. It automatically creates the query to perform these operations. To use Hibernate, it is required to create Java classes that represents the table in the database and then map the instance variable in the class with the columns in the database. The architecture is depicted in figure 8. Hibernate architecture has three main components:

- **Connection Management:** Hibernate Connection management service provides efficient management of the database connections. Database connection is the most expensive part of interacting with the database as it requires a lot of resources of open and close the database connection.
- **Transaction management :**Transaction management service provides the ability to the user to execute more than one database statements at a time.
- **Object relational mapping:** It is the technique of mapping the data representation from an object model to a relational data model. This part of Hibernate is used to select, insert, update and delete the records form the underlying table. When we pass an object to a Session.save() method, Hibernate reads the state of the variables of that object and executes the necessary query.

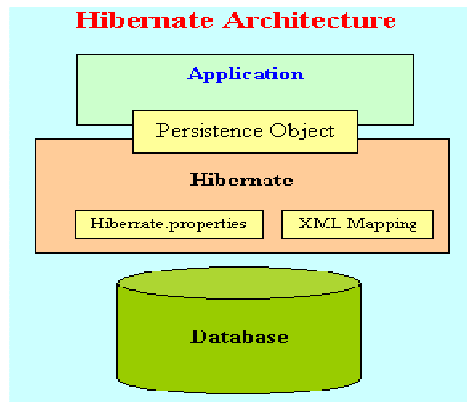


FIGURE 8: Hibernate Architecture.

7. CONCLUSION

HABIT is an off-line, writer independent handwriting analysis system that predicts the personality of a person. The input to the system is a scanned image of a handwriting sample. The output is a set of personality traits. To predict the actual personality of the individual there are various features, such as slant, size, pressure, upper zone loops, lower zone loops, word spacing, line spacing, page margins etc. HABIT utilizes the slant of letters, slant of the baseline, pen pressure and size of the letters as found in the person's handwriting sample. The implemented method here proves successful in analyzing handwriting irrespective of the language used. The analysis is completely devoid of user intervention. But stress has to be laid to the fact that the image acquisition process requires the aid of a graphologist. Enhancements possible on the implemented system include using the following features too in the traits prediction process.

- Position of 't' bar : Position of 't' bar indicates the self-esteem of a person. Using regression lines the distance between the 't' bar and bottom point or top point can be calculated. Space between letters:
- Perpendiculars are dropped from the end point of the first letter and starting point of the second letter. Length of the hypotenuse joining the two perpendiculars gives the distance between the two letters. Space between lines:
- Using Pythagoras theorem, distance between the perpendiculars can be calculated

8. REFERENCES

- 1 Shitala Prasad, Vivek Kumar Singh, Akshay Sapre "Handwriting Analysis based on Segmentation Method for Prediction of Human Personality using Support Vector Machine", International Journal of Computer Applications, Volume 8– No.12, October 2010.
- 2 Champa H N, K R AnandaKumar, "Automated Human Behavior Prediction through Handwriting Analysis", IEEE First International Conference on Integrated Intelligent Computing, 2010.
- 3 Sunday Olatunbosun, Aaron Dancygier, Jayson Diaz, Stacy Bryan, and Sung-Hyuk Cha,"Automating the Lewinson-Zubin Handwriting Personality Assessment Scales", Proceedings of Student-Faculty Research Day, CSIS, Pace University,2009.
- 4 Srihari S.N., Sung-Hyuk Cha and SangjikLee,"Establishing handwriting Individuality using pattern recognition techniques",Proceedings of the Sixth International Conference on Document Analysis and Recognition, 2001.

- 5 AmeerBensefia , Ali Nosary, Thierry Paquet, Laurent Heutte, "Writer Identification by Writer's Invariants", Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition", pp 274-279, 2002.
- 6 Srihari S.N., Sung-Hyuk Cha and Sangjik Lee, "Establishing handwriting Individuality using pattern recognition techniques", Proceedings of the Sixth International Conference on Document Analysis and Recognition, 2001.
- 7 N Mogharreban, S Rahimi, M Sabharwal, "A Combined Crisp and Fuzzy Approach for Handwriting Analysis", IEEE Annual Meeting of the Fuzzy Information, vol 1, pp 351-356, 2004.
- 8 Ondrej Rohlik, "Handwritten Text Analysis", Department of Computer Science and Engineering Laboratory of Intelligent Communication Systems, University of West Bohemia In Pilsen, 2001.

Brain Tumor Extraction from T1- Weighted MRI using Co-clustering and Level Set Methods

S.Satheesh

*Dept. of ECE,
G. Narayanamma Institute of Technology and Science,
Hyderabad, India.*

satheesh.s17@gmail.com

Dr.K.V.S.V.R Prasad

*Dept. of ECE,
D.M.S.S.V.H. College of Engineering,
Machilipatnam, India.*

kvsvr@yahoo.com

Dr.K.Jitender Reddy

*Dept. of Radiology and Imaging Sciences,
Apollo Health City,
Hyderabad, India.*

kjitenderreddymdrd@gmail.com

Abstract

The aim of the paper is to propose effective technique for tumor extraction from T1-weighted magnetic resonance brain images with combination of co-clustering and level set methods. The co-clustering is the effective region based segmentation technique for the brain tumor extraction but have a drawback at the boundary of tumors. While, the level set without re-initialization which is good edge based segmentation technique but have some drawbacks in providing initial contour. Therefore, in this paper the region based co-clustering and edge-based level set method are combined through initially extracting tumor using co-clustering and then providing the initial contour to level set method, which help in cancelling the drawbacks of co-clustering and level set method. The data set of five patients, where one slice is selected from each data set is used to analyze the performance of the proposed method. The quality metrics analysis of the proposed method is proved much better as compared to level set without re-initialization method.

Keywords: Magnetic Resonance Imaging, Tumor Extraction, Co-clustering Method, Level Set Method.

1. INTRODUCTION

Brain tumor is considered to be one of the significant diseases which require a controlled and timely diagnosis and treatment. It alarms an emergence of improved medical imaging techniques to be developed for treating the disease [1]. These technological innovations lead to early pathological diagnosis, their follow-ups, planning and guidance for surgeries along with quantitative analysis of the images. Among all imaging techniques, Magnetic Resonance Imaging (MRI) plays a vital role due to its ability in generating multi planar, good contrast, spatial resolution with anatomical details and capability to create 3D images for analyzing anatomy of brain in more in depth way in order to identify pathologies.

In brain tumor analysis, segmentation of abnormal tissues, anatomical structures and pathologies from MRI in particular, plays a predominant role. The results from this segmentation are the foundation for further analysis. It is necessary to change the segmentation methods depending on the hard and soft tissues and image modalities. In addition to, segmentation of MR brain images is a daunting task because they generally involve a large amount of data. While undergoing MRI

some artifacts will occur due to the patient's movements. Leading to the soft tissue boundaries which also can not be well defined.

When we deal with brain tumors, various other problems also arise making their segmentation more difficult. There are large number of types of tumors having a variety of shapes and sizes. These may be located at any place in the brain with various intensities. Sometimes, some of them will deform the surrounding structures that alter the intensities around the tumor. With the existence of different MR acquisition protocols different information is provided on the brain highlighting the region of the tumor.

Several segmentation methods such as thresholding, watershed, level set, zero crossing and region-based segmentation are used for tumor segmentation. In these methods, some methods are region based and they have a main disadvantage at the boundary of tumors. They suffer from misclassification of pixels and hence, it is hard to have a crisp region of tumor. Some of the methods are edge based and are suffer from initialization problems. But the medical images have both of two properties that are required to be overcome to get effective brain tumor extraction. Looking at the advantages of boundary based and region based methods, the third class of tumor segmentation methods was designed , which is the combination of region and boundary based techniques.

The first approach of this type was presented by Zhu and Yang [2].In their study they used thresholding and morphological operations. Law et al.,[3] proposed another method by integrating FCM clustering with the conventional snake. Chen et al.,[4] presented a new hybrid framework by integrating Gibbs model, marching cubes and parametric deformable models. Ho et al.,[5] presented a method to segment brain tumors by combining level set and fuzzy clustering. Taheri et al.,[6] combined threshold based method and level set to segment the brain tumor.

It is observed that co-clustering [7, 8], which is the region based segmentation method, is the effective clustering technique for the brain tumor extraction [9]. The level set without re-initialization [10], which is edge-based technique, is also efficient for the tumor extraction from MR brain images. Hence, in this paper both co-clustering region based technique is integrated with level set [10] edge based technique for taking advantages of both techniques while reducing their drawbacks to acquire effective tumor extraction of MR brain image. For this hybrid approach firstly tumor is extracted using co-clustering and then given this as the initial contour to level set. The experimental analysis of the proposed method is proved much better when compared with level set without re-initialization.

The overall structure of the research paper is demonstrated as follows; the second section provides the overview of co-clustering method. Third section illustrates level set without re-initialization overview. Fourth section describes the proposed technique steps, which is combining the co-clustering and level set methods. Fifth section illustrates regarding collected results with the evaluation of performance based on chosen evaluation metrics. Sixth section summarizes the proposed methodology of the research problem and gives future recommendations.

2. CO-CLUSTERING

The clustering is a collection of similar gray levels in the image, where gray levels are divided into diverse segments. On the other hand, co-clustering is partitioning of rows and columns simultaneously for an image [7, 8]. This kind of algorithm is significant in finding k-cluster of MR brain images. The T1-weighted MR brain images are skull stripped using [9] before applying to co-clustering algorithm.

The algorithm of co-clustering is summarized as follows:

Input: Skull stripped T1-weighted MR brain image of size $i \times j$ ($I_{i,j}$) and Number of clusters (k)

Output: Segmented MR brain image with k clusters

1. Form $I_n = D_1^{-1} \times I \times D_2^{-1}$ where $D_1(i, i) = \sum_j I_{i,j}$ and $D_2(j, j) = \sum_i I_{i,j}$
2. Compute $L = \lceil \log_2 k \rceil$ singular vectors of I_n
3. Apply singular value decomposition (SVD) technique on I_n to obtain
 $[U \ S \ V] = \text{SVD}(I_n)$
 where U and V represent the left and right eigenvector of 2^{nd} to $(L+1)^{\text{th}}$ eigen values.
 $U = [u_2, u_3, \dots, u_{L+1}]$ and $V = [v_2, v_3, \dots, v_{L+1}]$
4. Form the matrix $Z = \begin{bmatrix} D_1^{-1} \times U \\ D_2^{-1} \times V \end{bmatrix}$
5. The k-means algorithm is applied on the L -dimensional data Z to obtain the k number of clusters.
6. Find out the mean of the centers and their indices from the obtained clusters.
7. Extract the segmented portion by indexing the obtained L -dimensional data with respective to original image.
8. Apply morphological region filling operator to refine the tumor region.

Co-clustering is found taking advantage from the duality between the columns and rows, in order to deal with the high dimensional data influentially. There is a limitation of over segmentation, associated with co-clustering based image segmentation. Thus, in order to overcome the issue of over segmentation, integration of edge based segmentation will give better results.

3. LEVEL SET WITHOUT RE-INITIALIZATION

Osher and Sethian [11] initially introduced level set methods for capturing moving fronts. The level set method is the effective way to demonstrate active contour, which helps in MR brain tumor segmentation. Recently, several research works have been done on the geometric active contours [12-16], where active contours applied through level set method to address the broad range of image segmentation problems in image processing and computer vision.

Active contours employed through level set methods can be formulated as zero level set of a time dependent function ϕ that varies according to the equation (1).

$$\frac{\partial \phi}{\partial t} + F|\nabla \phi| = 0 \tag{1}$$

Equation (1) is known as level set equation. Here, F is called speed function depending on image data and level set function ϕ . While implementing the level set method, it is compulsory to keep the evolving level set function close to a signed distance function in order to sustain stable curve evolution. Yet, the process of re-initialization makes the total computation expensive. It also causes numerical error in the location of the zero level set.

In this paper, we have used a level set evolution method which is based on energy penalty term without re-initialization introduced by Li Chunming et al., [10].

Let I be an image, and g be the edge indicator function defined by equation (2).

$$g = \frac{1}{1 + |\nabla G_{\sigma} * I|^2} \tag{2}$$

where G_σ is the Gaussian kernel with standard deviation σ .

The external energy for a function $\phi(x, y)$ is defined as

$$\mathcal{E}_{g,\lambda,v}(\phi) = \lambda \mathcal{L}_g(\phi) + v \mathcal{A}_g(\phi) \quad (3)$$

where $\lambda > 0$ and v are constants.

The terms $\mathcal{L}_g(\phi)$ and $\mathcal{A}_g(\phi)$ can be defined as

$$\mathcal{L}_g(\phi) = \int_{\Omega} g \delta(\phi) |\nabla \phi| dx dy \quad (4)$$

$$\mathcal{A}_g(\phi) = \int_{\Omega} g H(-\phi) dx dy \quad (5)$$

where δ is the univariate Dirac function, and H is the Heaviside function.

The total energy function is defined as

$$\mathcal{E}(\phi) = \mu \mathcal{P}(\phi) + \mathcal{E}_{g,\lambda,v}(\phi) \quad (6)$$

Here, the total energy function contains both an internal energy term and an external energy term. The internal energy term $\mathcal{P}(\phi)$ penalizes the deviation of the level set function from a signed distance function and the external energy term $\mathcal{E}_{g,\lambda,v}(\phi)$ drives to motion of the zero level set to the required image features like object boundaries.

The evolution equation of the level set function is defined as

$$\frac{\partial \phi}{\partial t} = \mu \left[\Delta \phi - \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right] + \lambda \delta(\phi) \text{div} \left(g \frac{\nabla \phi}{|\nabla \phi|} \right) + v g \delta(\phi) \quad (7)$$

4. PROPOSED METHODOLOGY

The hybrid approach is being proposed to acquire the best possible methodology for effective tumor extraction results. The proposed algorithm can be summarized as follows:

Step 1: The first step of the combined methodology is to read the T1-weighted MR brain image.

Step 2: Apply morphological operations [9] on T1-weighted MR brain image to remove non brain data (skull, fat, skin, muscle).

Step 3: After removing the non brain region from the T1-weighted MR brain image, use co-clustering algorithm, in order to extract the tumor region [9].

Step 4: The tumor region obtained in the step 3, is to be defined as initial contour for level set method.

Step 5: Use the initial contour, which is defined in the step 4 of the methodology in order to obtain final tumor contour, by level set without re-initialization [10], which is defined in equation (7).

In order to understand the proposed methodology more clearly diagrammatic illustration is provided in Figure 1. It can be observed that after doing initial segmentation of tumor through co-clustering, there is a need to refine the segmentation by the level set without re-initialization. Thus, after applying both the methods in this proposed methodology, evaluation of the segmentation is done systematically.

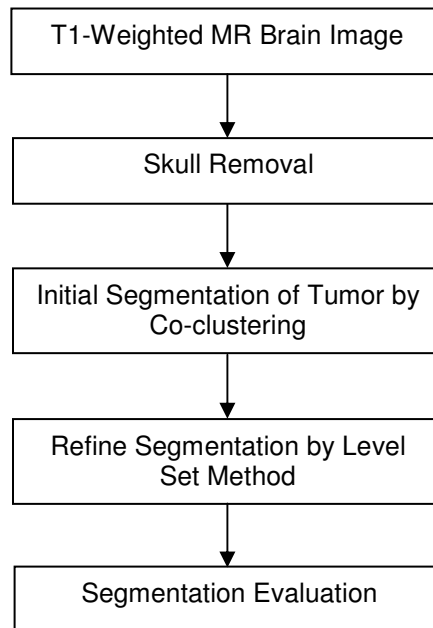


FIGURE 1: Diagrammatic Illustration of the Proposed Methodology.

5. RESULTS AND DISCUSSION

5.1 Data Sets

Data sets were collected from the Department of Radiology and Imaging Science, Apollo Health City, Hyderabad, India and Lucid Diagnostics, Hyderabad, India. These data sets have been acquired on 1.5T Philips achieva apparatus and 1.5T G.E apparatus using an axial T1-weighted sequence with contrast agent. The proposed method was verified on MR brain image data sets of five patients named as Patient 1 to patient 5 where one slice was selected from the data set of each patient to analyze the performance of the proposed method.

5.2 Evaluation Metrics

The evaluation metrics for analyzing the proposed methodology includes:

Similarity Index (SI): SI is the measurement, which provides true-segmented region relative to the total segmented region.

$$SI = \frac{2TP}{2TP+FP+FN} \times 100\% \quad (8)$$

where TP is the number of pixels detected correctly, FP is the number of pixels detected falsely as tumor and FN is the number of pixels detected falsely as non-tumor.

Correct Detection Ratio (CDR): The CDR value indicates the degree of trueness of the actual tumor.

$$CDR = \frac{TP}{TP+FN} \times 100\% \quad (9)$$

Total Segmentation Error (TSE): It is the sum of Under Segmentation Error (USE) and Over Segmentation Error (OSE).

$$TSE = USE + OSE \quad (10)$$

where $USE = \frac{FP}{TP+FN} \times 100\%$ and $OSE = \frac{FN}{TP+FN} \times 100\%$

5.3 Discussion

In this paper, we have evaluated the tumor extraction results based on the suitable evaluation metrics like SI, CDR and TSE, and analyzed the outcomes of the proposed methodology quality metrics values with the level set method [10]. The extracted results of proposed methodology of the tumor are demonstrated in Figure 2 for the slice 95 of patient 2. It is observed that close proximity to the manually segmented images by the experts and are better than level set method. The quantitative results attained by the proposed method in comparison with level set method are provided in Table 1.

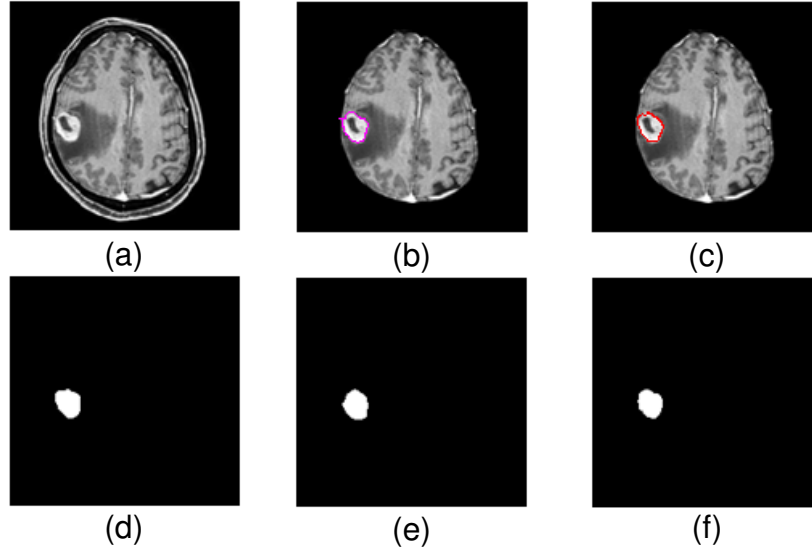


Figure 2: Tumor extraction result of patient 2 of slice 95 (a) One axial slice of the selected tumor class (b) Initial contour for proposed method (c) Final contour of proposed method (d) Final extracted tumor by the proposed method (e) Extracted tumor by the level set without re-initialization (f) Manually segmented tumor.

Patient	Slice No.	Method	SI (%)	CDR (%)	TSE (%)
1	101	Level Set	85.246	100	34.615
		Proposed Method	88.954	96.795	24.038
2	95	Level Set	91.12	100	19.491
		Proposed Method	96.305	98.488	07.5577
3	167	Level Set	80	100	50
		Proposed Method	87.826	91.818	25.455
4	120	Level Set	83.119	100	40.618
		Proposed Method	93.268	98.284	14.188
5	83	Level Set	81.569	100	45.192
		Proposed Method	90.942	97.5	19.423

Table 1: Comparison of evaluation metrics obtained using the proposed method and level set method.

It is observed from the Table 1 that the SI of the proposed methodology varies from 87.826% to 96.305% but for level set method it is 80% to 91.12%. The CDR ranges from 91.818% to

98.488% for the proposed method. But for the level set method for all the five images CDR is 100% which is due to under segmentation. The TSE ranges from 07.5577% to 25.455% for the proposed method and for level set method it is 19.491% to 50%, which confirms good results of the proposed methodology as compared to level set without re-initialization.

6. CONCLUSION

The proposed tumor extraction method was tested on five abnormal brain slices of different patients ranges from patient 1 to patient 5 and the performance was evaluated based on the SI , CDR and TSE evaluation metrics. It is observed that combining level set without re-initialization with the co-clustering technique in the proposed methodology reduces the segmentation error and provided much better quality metrics values as compared to the level set without re-initialization. In the future research, the effect of the prior information on the object boundary extraction with the level set method such as shape and size can be further analyze. Moreover, the performance of the image segmentation method can be evaluated with other quality metrics along with SI, CDR and TSE to analyze the results more efficiently.

Acknowledgement

We wish to express our sincere thanks to Dept. of Radiology & Imaging Sciences, Apollo Health City, Hyderabad, India and Lucid Diagnostics, Hyderabad, India for providing us with different MR Image datasets.

7. REFERENCES

- 1 Chen, Yunjie; Zhang, Jianwei; Pheng-Ann Heng; Xia, Deshen, "Chinese Visible Human Brain Image Segmentation," Image and Signal Processing, 2008. CISP '08. Congress on , vol.3, no., pp.639,643, 27-30 May 2008.
- 2 Yan Zhu; Hong Yan, "Computerized tumor boundary detection using a Hopfield neural network," Medical Imaging, IEEE Transactions on , vol.16, no.1, pp.55,67, Feb. 1997.
- 3 Law, A.K.W.; Hui Zhu; Chan, B.C.B.; lu, P. P.; Lam, F. K.; Chan, F.H.Y., "Semi-automatic tumor boundary detection in MR image sequences," Intelligent Multimedia, Video and Speech Processing, 2001. Proceedings of 2001 International Symposium on , vol., no., pp.28,31, 2001.
- 4 Chen Ting, Dimitris Metaxas., "A hybrid framework for 3D medical image segmentation," Medical Image Analysis, Volume 9, Issue 6, Pages 547-565, December 2005.
- 5 Ho, S.; Bullitt, E.; Gerig, G., "Level-set evolution with region competition: automatic 3-D segmentation of brain tumors," Pattern Recognition, 2002. Proceedings. 16th International Conference on , vol.1, no., pp.532,535 vol.1, 2002.
- 6 Taheri, S.; Sim Heng Ong; Chong, V., "Threshold-based 3D Tumor Segmentation using Level Set (TSL)," Applications of Computer Vision, 2007. WACV '07. IEEE Workshop on , vol., no., pp.45,45, Feb. 2007.
- 7 Dhillon, Inderjit S "Co-clustering documents and words using bipartite spectral graph partitioning". Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining -KDD '01, San Francisco, California,pages: 269--274 ,2001.
- 8 Rege, M.; Ming Dong; Fotouhi, F., "Co-clustering Documents and Words Using Bipartite Isoperimetric Graph Partitioning," Data Mining, 2006. ICDM '06. Sixth International Conference on , vol., no., pp.532,541, 18-22 Dec. 2006.

- 9 S.Satheesh, Dr.K.V.S.V.R Prasad, Dr.K.Jitender Reddy; "Automatic tumor extraction for contrast enhanced axial T1-weighted magnetic resonance brain images integrating co-clustering and morphological operations", International Conference on Signal, Image and Video Processing (ICSIVP), IITPatna , pp.214-218,January -2012.
- 10 Chunming Li; Chenyang Xu; Changfeng Gui; Fox, M.D., "Level set evolution without re-initialization: a new variational formulation," Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on , vol.1, no., pp.430,436 vol. 1, 20-25 June 2005.
- 11 Osher Stanley, Sethian James A., "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," Journal of Computational Physics, Volume 79, Issue 1,pp. 12-49, , November 1988.
- 12 Malladi, R.; Sethian, J.A.; Vemuri, B.C., "Shape modeling with front propagation: a level set approach," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol.17, no.2, pp.158,175, Feb 1995.
- 13 Hongyu Lu; Youming Yu; Shanglian Bao, "Note: On modeling techniques in active contours," Signal Processing (ICSP), 2012 IEEE 11th International Conference on , vol.2, no., pp.956,961, 21-25 Oct. 2012.
- 14 Li, Danyi; Li, Weifeng; Liao, Qingmin, "Active contours driven by local probability distributions," Image and Signal Processing (CISP), 2012 5th International Congress on , vol., no., pp.634,638, 16-18 Oct. 2012.
- 15 Si Yong Yeo, "Implicit active contours for N-dimensional biomedical image segmentation," Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on , vol., no., pp.2855,2860, 14-17 Oct. 2012.
- 16 Verma, N.; Muralidhar, G.S.; Bovik, A.C.; Cowperthwaite, M.C.; Markey, M.K., "Model-driven, probabilistic level set based segmentation of magnetic resonance images of the brain," Engineering in Medicine and Biology Society,EMBC, 2011 Annual International Conference of the IEEE , vol., no., pp.2821,2824, Aug. 30 2011-Sept. 3 2011.

INSTRUCTIONS TO CONTRIBUTORS

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

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

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

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with volume 7, 2013, IJIP is appearing with more focused issues. Besides normal publications, IJIP intends to organize special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

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

LIST OF TOPICS

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

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

- Storage and transmission

- Video coding algorithms and technologies for ATM/p

CALL FOR PAPERS

Volume: 7 - Issue: 5

i. Paper Submission: July 30, 2013

ii. Author Notification: September 15, 2013

iii. Issue Publication: October 2013

CONTACT INFORMATION

Computer Science Journals Sdn Bhd

B-5-8 Plaza Mont Kiara, Mont Kiara

50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6207 1607

006 03 2782 6991

Fax: 006 03 6207 1697

Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2013
COMPUTER SCIENCE JOURNALS SDN BHD
B-5-8 PLAZA MONT KIARA
MONT KIARA
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6207 1607
006 03 2782 6991

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