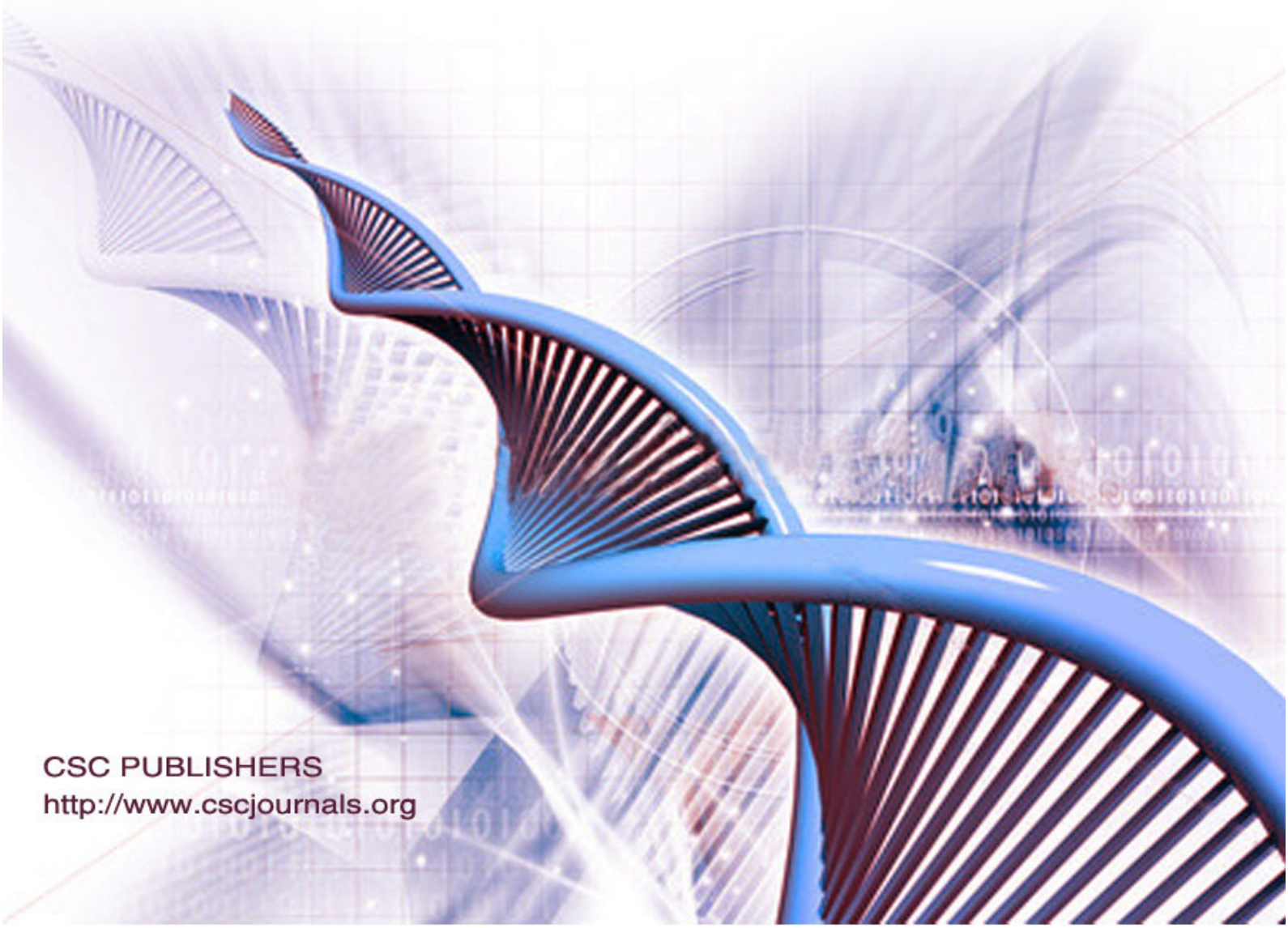
Volume 7 ▪ Issue 1 ▪ April 2013

INTERNATIONAL JOURNAL OF

BIOMETRICS AND BIOINFORMATICS (IJBB)

ISSN : 1985-2347
Publication Frequency: 6 Issues / Year



CSC PUBLISHERS
http://www.cscjournals.org

INTERNATIONAL JOURNAL OF BIOMETRICS AND BIOINFORMATICS (IJBB)

VOLUME 7, ISSUE 1, 2013

EDITED BY
DR. NABEEL TAHIR

ISSN (Online): 1985-2347

International Journal of Biometrics and Bioinformatics (IJBB) 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.

IJBB Journal is a part of CSC Publishers

Computer Science Journals

http://www.cscjournals.org

INTERNATIONAL JOURNAL OF BIOMETRICS AND BIOINFORMATICS (IJBB)

Book: Volume 7, Issue 1, June 2013
Publishing Date: 30-06-2013
ISSN (Online): 1985-2347

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

Published in Malaysia

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

CSC Publishers, 2013

EDITORIAL PREFACE

This is the *first* Issue of Volume *seven* of International Journal of Biometric and Bioinformatics (IJBB). The Journal is published bi-monthly, with papers being peer reviewed to high international standards. The International Journal of Biometric and Bioinformatics is not limited to a specific aspect of Biology but it is devoted to the publication of high quality papers on all division of Bio in general. IJBB intends to disseminate knowledge in the various disciplines of the Biometric field from theoretical, practical and analytical research to physical implications and theoretical or quantitative discussion intended for academic and industrial progress. In order to position IJBB as one of the good journal on Bio-sciences, a group of highly valuable scholars are serving on the editorial board. The International Editorial Board ensures that significant developments in Biometrics from around the world are reflected in the Journal. Some important topics covers by journal are Bio-grid, biomedical image processing (fusion), Computational structural biology, Molecular sequence analysis, Genetic algorithms etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 7, 2013, IJBB appears with more focused issues related to biometrics and bioinformatics studies. Besides normal publications, IJBB intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

The coverage of the journal includes all new theoretical and experimental findings in the fields of Biometrics which enhance the knowledge of scientist, industrials, researchers and all those persons who are coupled with Bioscience field. IJBB objective is to publish articles that are not only technically proficient but also contains information and ideas of fresh interest for International readership. IJBB aims to handle submissions courteously and promptly. IJBB objectives are to promote and extend the use of all methods in the principal disciplines of Bioscience.

IJBB 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 IJBB. 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. IJBB provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members
International Journal of Biometric and Bioinformatics (IJBB)

EDITORIAL BOARD

Dr Shunzhou Yu
General Motors Global R&D Center
United States of America

Dr Christopher Taylor
University of New Orleans
United States of America

Dr Horacio Pérez-Sánchez
University of Murcia
Spain

TABLE OF CONTENTS

Volume 7, Issue 1, June 2013

Pages

Knee Joint Articular Cartilage Segmentation using Radial Search Method, Visualization and Quantification

M. S. Mallikarjuna Swamy *ms_muttad@yahoo.co.in*
Department of Biomedical Engineering and Research Centre
Bapuji Institute of Engineering and Technology
Davangere-577004, Karnataka, India

Mallikarjun S. Holi *msholi@yahoo.com*
Department of Biomedical Engineering and Research Centre
Bapuji Institute of Engineering and Technology
Davangere-577004, Karnataka, India

Abstract

Knee is a complex and highly stressed joint of the human body. Articular cartilage is a smooth hyaline spongy material between the tibia and femur bones of knee joint. Cartilage morphology change is an important biomarker for the progression of osteoarthritis (OA). Magnetic Resonance Imaging (MRI) is the modality widely used to image the knee joint because of its hazard free and soft tissue contrast. Cartilage thickness measurement and visualization is useful for early detection and progression of the disease in case of OA affected patients. In the present work, knee joint MR images of normal and OA affected are processed for segmentation and visualization of cartilage using semiautomatic method. The radial search method is used with minor modifications in search area to reduce computation time. Cartilage thickness and volume is measured in lateral, medial and patellar regions of femur. The overall accuracy of measurements is determined by comparing the measurements with another semiautomatic method based on edge detection and interpolation. It is observed a good correlation between quantification of cartilage in two methods. The method takes less time for segmentation because of reduced manual steps. The reduced cartilage thickness and volume is observed in OA affected knee of different level of progression.

Keywords: Cartilage, Image Segmentation, Knee Joint, MRI, Osteoarthritis.

1. INTRODUCTION

The knee joint is the largest and most complex synovial joint of the human body. It is a major weight bearing joint of the body made up of condyles of femur, condyles of tibia and posterior surface of patella. Articular cartilage is a thin layer between the femur and tibia bones. It is a soft tissue at the end of bones that allows the joint to move freely. The knee joint contains a small amount of fluid in a cavity that lubricates the cartilage called synovial fluid. Osteoarthritis is a common disease of the knee joint affecting the elderly people. It occurs when cartilage becomes soft and gets eroded due to continuous wear and tear movements and with ageing. The OA affected knee joint often leads to inflammation, decrease in motion of joint due to stiffness, and formation of bone spurs (tiny growths of new bone). This decreases the ability of the cartilage to work as a shock absorber to reduce the impact of stress on the joints. The remaining cartilage wears down faster, and eventually, the cartilage in some regions may disappear altogether, leaving the bones to rub against one another during motion leading to formation of bone spurs. With OA, synovial fluid does not provide proper lubrication, which leads pain, inflammation and restriction of movements at the joints. There is no artificial material that can replace only the cartilage at the joint. In a clinical assessment study conducted on Indian population consisting of 362 elderly of more than 65 years, osteoarthritis was present in only 50.2% of the elderly aged 65-74 years, whereas it was 97.7% in elderly aged 84 years and above [1]. OA affected nearly 27

million Americans according to the study in 2007 [2]. After the age of 50, women are more prone OA than men [3].

1.1 MR Imaging of knee joints and clinical findings

MRI is non-invasive and repetitive imaging study of an individual is possible without side effects. The assessment of cartilage dimensions is important for the study of the progression of cartilage damage due to OA. MRI images are widely used for diagnosis of knee joint abnormalities. This imaging modality provides in-vivo and in-vitro information of anatomical structures. MRI can visualize cartilage, bone and other surrounding soft tissues distinctly. A comparison of MR evaluation of the morphology of articular cartilage with data from histology shows good correlation across those two modalities [4]. For quantification of cartilage thickness, volume and progressive assessments image processing techniques are used. Association between clinical features and MR image findings of knee joint are evaluated and it is found that a large joint effusion is associated with pain and stiffness [5]. The presence of osteophytes in the patella femoral compartment is also associated with pain. All other abnormalities in cartilage, menisci and subchondral cysts can be found in MR imaging only. The use of MRI for diagnosis and assessment of cartilage defect repairs has been studied [6]. MR imaging protocols like fat suppressed spoiled gradient echo sequence and the fast spin echo sequence are accurate and reliable for evaluating surface defects of articular cartilage. MR imaging findings in different stages of disease and correlation with clinical findings is studied [7]. Cartilage lesions, bone marrow edema pattern and meniscal lesions are well detected on MR images in patients with advanced OA. Anatomical variants in the knee are frequent findings on MRI. Thorough knowledge and familiarity with variant and its pathological nature are important for accurate interpretation of imaging studies [8].

1.2 Knee cartilage segmentation and quantification methods

Knee joint image segmentation is a very challenging task because of its complexity. Segmentation methods for knee joint can be classified into three categories based on manual intervention required, namely manual, semiautomatic and fully automatic. The manual segmentation methods are laborious and time consuming [9]. Semiautomatic methods are developed to reduce the manual intervention by automating few steps of processing. Fully automatic methods involve advanced and complex processing steps with certain limitations. Zohara et al. [10] developed a semiautomatic method, initially cartilage is segmented manually by marking the consecutive points along the articular contour curves with a typical spacing of 0.5-1.0 mm. An interpolated cubic B-spline curve is fitted for these points. Cashman et al. [9] developed an algorithm using edge detection and thresholding. Boundary discontinuities are bridged using B-spline interpolation and recursive region growing procedure is used in segmentation of bone. In radial search method developed by Poh et al. [11], a threshold method is used to detect the inner boundaries along the radial lines. The algorithm searches the boundary of cartilage from approximate center of femur bone region. In the method developed by Gamio et al. [12] Bezier splines are used. The control points are placed inside the cartilage following its shape to create a Bezier spline. Rays perpendicular to the spline on the control points are traced to find the bone cartilage interface. The edges are found based on the first derivative of brightness using bicubic interpolation along the line profiles. In the graph cut method developed by Shim et al. [13] seeds are placed manually (curvilinear marks) over specific anatomic regions. The seeds are propagated to neighboring pixels and then segmented. A fully automatic method using voxel classification is developed by Jenny et al. [14]. The algorithm is based on kNN classifier to reduce processing time. In the 2-D active contour algorithm developed by Claude et al. [15], a local coordinate system (LCS) is developed for the femoral and tibial cartilage boundaries for the measurement of thickness and volume. Tang et al. [16] proposed segmentation of articular cartilage surfaces using snakes, and a gradient vector flow (GVF) based external force. GVF snake is made more stable and converge to the correct surfaces, directional gradient is used to produce the gradient vector flow. Segmentation method is developed for multiple interacting surfaces belonging to multiple interacting objects, called LOGISMOS (layered optimal graph image segmentation of multiple objects and surfaces) by Yin et al. [17]. The approach is based on the algorithmic incorporation of multiple spatial inter relationships in a single n-dimensional graph,

and followed by graph optimization. Dodin et al. [18] segmented the cartilage by resampling the MRI in the neighborhood of the bone surface. Texture analysis techniques are optimized by filtering and then cartilage is discriminated as a bright and homogeneous tissue. This excludes soft tissues and enables the detection of the external boundary of the cartilage. A Bayesian decision criterion is used for automatic separation of the cartilage and synovial fluid. An unsupervised method is developed for segmentation and quantification of knee features by Tamez Pena et al [19].

Even though there are many algorithms developed for knee joint cartilage segmentation still there is a scope for better segmentation and precise quantification of cartilage because of the complexity involved in segmentation algorithms. In this work, a semiautomatic technique is developed to segment cartilage from knee MRI with reduced computation time and good accuracy in measurements.

2. METHODS

The knee joint MR images are obtained from JSS Medical College and Hospital, Mysore which includes normal and OA images. The MR images are obtained using Siemens 1.5T MRI system in fat suppressed spoiled gradient recalled (SPGR) image protocol. The imaging parameters for the sequence are: TR/TE: 16.3/4.7 ms, matrix: 256x256, FOV: 140 mm, slice thickness: 0.7 mm, x/y resolution: 0.365/0.365 mm. For this study approval is obtained from the hospital ethics committee to review the medical records and images of the patients who had been clinically diagnosed for knee joint problems and undergone MR imaging. The clinical sample of this study consists of 55 knee joint MRI dataset corresponding to one knee of each of 55 individuals. The study population ranged from 18 to 75 years. The MR images are processed using matlab 7.1 software for segmentation and visualization of cartilage.

The input knee MR images are preprocessed for noise removal using median filter. Median filter eliminates noise and preserves the edges in images. The cartilage is segmented using radial search method described by Poh et al. [11] with modification in search area. The processing steps for cartilage segmentation, thickness measurement, volume computations and 3D visualization are shown in Fig. 1.

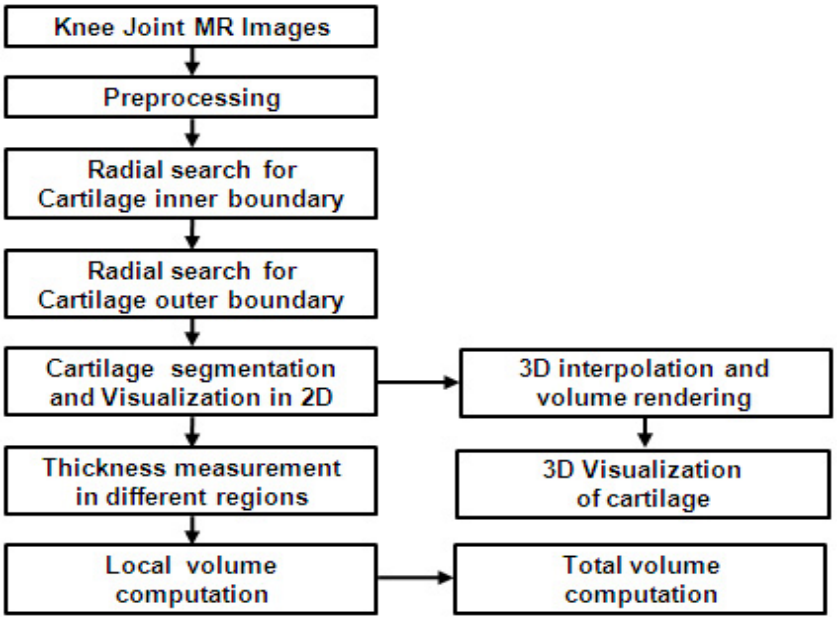


FIGURE 1: Processing steps of segmentation, thickness measurement, volume measurement and 3D visualization.

The sample points on inner boundary and outer boundary are obtained using radial search method. Using inner and outer boundary sample points a mask is developed for the cartilage. Then the cartilage is segmented from the knee MRI. The segmented cartilage thickness is computed along a normal to its inner boundary. The 2D thickness maps of cartilage in MRI sequence are saved. The voxels in between the MR slices are obtained using 3D interpolation. The articular cartilage is volume rendered using 3D texture mapping technique. The volume of the cartilage is computed by cumulatively adding the local volumes obtained.

2.1 Detection of cartilage boundary

Based on the average intensity of pixels in the cartilage region a threshold value is selected. Approximate center of the femur is selected as an origin for radii search. The search starts from this origin along a radial line. The length of the radial line is chosen to reach beyond the outer boundary. Search procedure is initiated to detect first pixel which belongs to articular cartilage along the radial line. In the inner circle of radius *r* there is no possibility of detecting the inner boundary. Therefore the search is started after *r* number of pixels away from the origin. A change in threshold level indicates the inner boundary point called b_{i1}. The coordinates of point b_{i1} are saved as inner cartilage boundary sample point. The search is continued along the next radial line which is θ° away from the previous radial line. This search results in another boundary point along the inner cartilage called b_{i2}. The search is repeated for *n* number of radial lines for an angular increment of θ° each time. The search is conducted for θ° increment from 0 to 300° which covers the different regions of cartilage. The possible boundary sample points along the inner boundary are obtained. The search is conducted for increment of 5°. This results in nearly 60 boundary sample points along the femur cartilage interface. The sample points on the boundary are further increased by interpolation. Fig. 2 depicts the radial search algorithm.

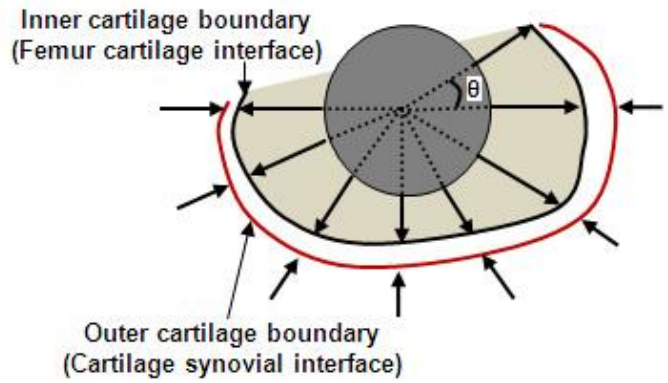


FIGURE 2: Radial search method.

Let, (x_0, y_0) is the origin of search. The search starts at r_k distance (pixels) away along the radial line from the origin. The pixel searched along a radial line is *0* to *K-1*. Search is conducted in *N* number of radial lines. The coordinate point (x_k, y_k) at which the threshold is detected is saved as boundary point. The equations of x_k and y_k are given as

$$x_k(n) = x_0 + r_k \cos(\theta_n) \tag{1}$$

$$y_k(n) = y_0 + r_k \sin(\theta_n) \tag{2}$$

where $k = 0,1.....,K-1$ and $n = 0,1,....,N$

For outer boundary detection the search is restarted along the radial line r_0 pixels away from the origin. The search is towards the origin from outside of the outer boundary. The pixel belongs to the cartilage along the radial line is searched to detect the threshold change. When the threshold level is detected, the coordinates of the pixel are saved as outer boundary point b_{o1}. The search is repeated for the radial lines drawn from 0° to 300° with an angular increment of 5°. This results in nearly 60 points along the outer boundary.

2.2 Segmentation of articular cartilage

The inner boundary sample points and outer boundary sample points of cartilage are obtained using above mentioned procedure. Fig. 3 shows the images at different steps of radial search algorithm and cartilage segmentation.

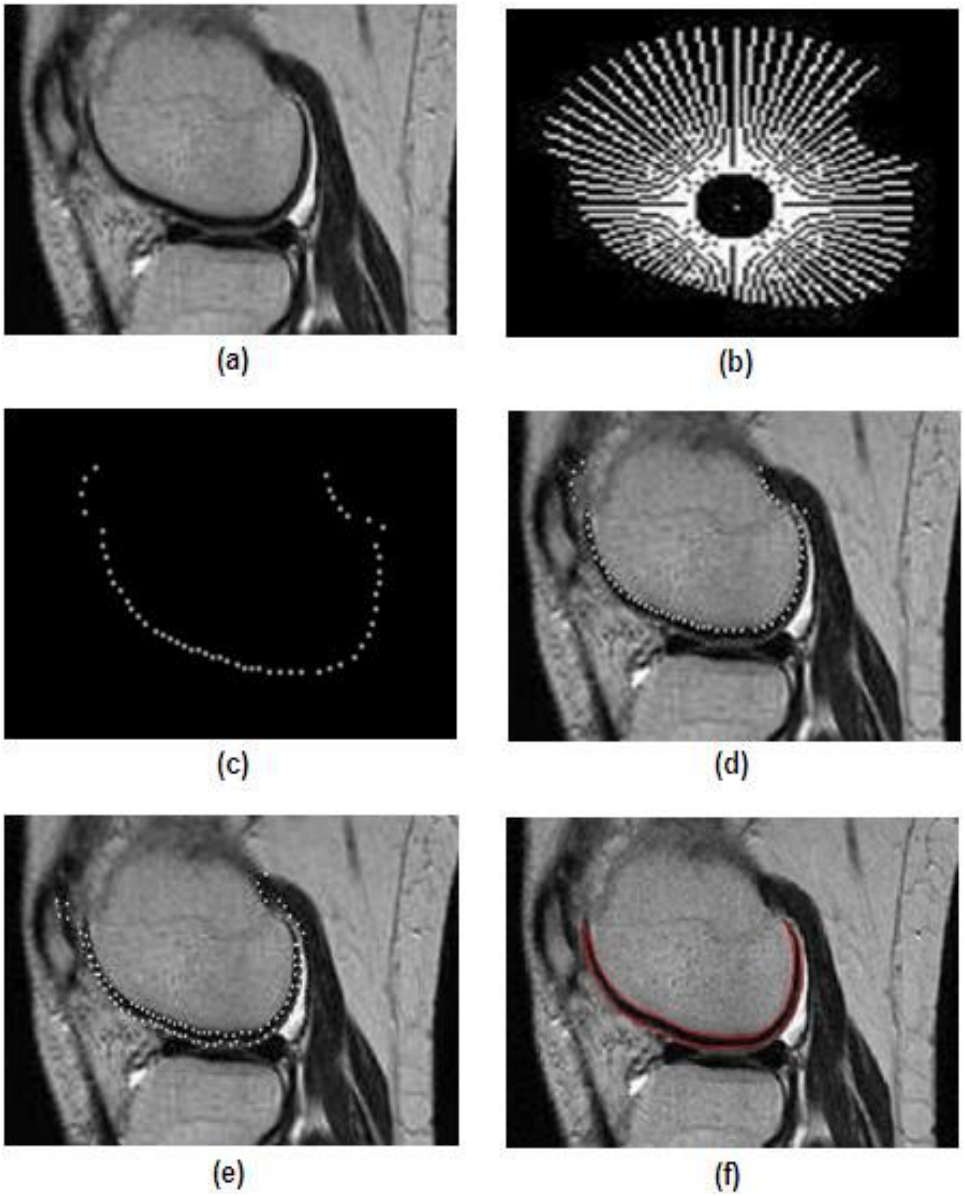


FIGURE 3: Images at different steps of processing (a) knee MR image (b) radial search (c) detection of boundary points (d) detected inner boundary points overlapped on input image (e) detected boundary points overlapped on input image (f) segmented cartilage overlapped on input image

A mask is generated using the boundary sample points. The articular cartilage is segmented from the MRI of knee joint. The manual intervention is required only to initiate the radial search process by selecting the origin of search for each image. Other processing steps are automatic.

2.3 Thickness and volume measurement of articular cartilage

The thickness of cartilage is measured in lateral, medial and patellar regions. The cartilage thickness is computed along the local normal of measurement point on the inner boundary towards the outer boundary of the cartilage. Euclidian distance between the inner boundary point b_i (x_1, y_1) to outer boundary point b_o (x_2, y_2) along the normal is computed using

$$t = \sqrt{\left(\left(x_1 - x_2 \right)^2 + \left(y_1 - y_2 \right)^2 \right)} \tag{3}$$

This results in n number of cartilage thickness (t_i) values in a region. The arithmetic mean of thickness (t_m) in a region is computed as

$$t_m = \frac{\sum t_i}{n} \tag{4}$$

The mean thickness value cartilage of a region from consecutive MRI slices is computed and average value of thickness of cartilage (t_a) in a region is calculated. The standard deviation (SD) of mean thickness from its average is computed as

$$SD = \sqrt{\frac{\sum (t_m - t_a)}{n-1}} \tag{5}$$

The coefficient of variance (COV) of cartilage thickness is computed as

$$COV = \frac{SD}{t_a} \times 100 \tag{6}$$

The local volume of cartilage is computed region wise. The total volume of cartilage in a region is summation of local volumes obtained of different regions. Volume measurement of cartilage is repeated in the same region. Average of cartilage volume (V_a) and SD is calculated. Let the total volumes in lateral, medial and patellar regions be V_l, V_m and V_p respectively. The total volume (V_T) of entire femur cartilage of knee is computed adding the volumes of different regions.

$$V_T = V_l + V_m + V_p \tag{7}$$

The segmented cartilage from MRI slice is saved as 256 x 256 size image along with its 2D texture. This is repeated for all the slices to create a stack of images of segmented cartilage. A 3D array of size (256 x 256 x 1x n) is created. The segmented cartilage images are saved into 3D array. The 3D data is volume rendered using 3D texture mapping technique. The texture is interpolated on the entire volume.

To compute the accuracy of thickness and volume of articular cartilage the measured values are compared with standard values obtained with another semiautomatic method [20] based on canny edge detection, manual marking of boundary points and B-spline interpolation. The thickness and volume of cartilage is quantified using validated semiautomatic method for the entire knee joint MRI data set. The measured thickness and volume using radial search method

are compared with the readings of the method as actual. The overall accuracy of measurements is determined using RMS (root mean square) residual difference and RMS mean error.

3. RESULTS

MR images of 55 cases including 7 normal and 48 OA affected knee joint subjects are obtained for this study. The population includes 27 male, 28 female with varying age group of 18 to 75. After computing cartilage thickness of these cases, the OA cases are separated from normal cases. The average thickness of cartilage in different regions of cartilage is measured from MRI data set of different subjects. The average cartilage thickness, SD and COV of different regions of femur of 2 normal and 4 OA affected subjects of different degree of disease severity are tabulated in Table 1.

Clinical symptom	Articular Cartilage Thickness in mm								
	Lateral			Medial			Patella		
	Avg	SD	COV	Avg	SD	COV	Avg	SD	COV
Normal1	2.08	0.06	3.54	2.04	0.05	2.75	2.15	0.13	4.86
Normal2	2.09	0.07	3.78	2.07	0.04	2.81	2.12	0.1	4.75
OA1	1.99	0.12	6.95	1.97	0.19	3.06	1.99	0.08	3.52
OA2	2.01	0.14	6.28	1.93	0.11	6.12	1.96	0.14	6.89
OA3	1.89	0.12	6.11	1.89	0.18	7.38	1.97	0.13	5.41
OA4	1.96	0.16	9.67	1.91	0.19	8.45	1.88	0.23	10.97

TABLE 1: Measured thickness of articular cartilage.

The local volumes of cartilage at different regions are computed for the MRI data set of a subject. The volume measurement of a region is repeated 3 times and average of volume measurement is computed along with SD. The Table 2 shows the computed volume of cartilage and SD in different regions for the same subjects of Table 1.

Clinical symptom	Articular cartilage volume in mm^3						Total volume ±SD
	Lateral		Medial		Patellar		
	Mean	SD	Mean	SD	Mean	SD	
Normal1	2105	118.86	2138	101.95	1173	981.6	5406
Normal2	2153	128.37	2189	107.96	1162	917.6	5491
OA1	2024	111.34	2074	116.18	1169	986.7	5282
OA2	2013	105.23	2027	129.51	1131	1038.5	5161
OA3	1996	139.68	1979	174.94	1113	1061.63	5102
OA4	1911	234.58	1925	218.03	998	1176.56	4801

TABLE 2: Measured volume of articular cartilage.

The total cartilage volume is obtained adding the volumes of cartilage of different regions. The volume is computed for normal and OA knee joints at different levels of progression. The Table 3 shows the accuracy of measurement of radial search method in comparison with measurements obtained using validated edge based method which is considered as reference. The accuracy is determined by root mean square (RMS) residual difference in measurements and RMS mean error. The measurements of cartilage at each point are compared with standard measured value and then root of the squared difference of individual measurements is computed. Percentage error of measurement is also computed in each case.

Clinical symptom	Accuracy of cartilage thickness measurement in mm								
	Lateral			Medial			Patella		
	Residual	Mean Error	Percent Error	Residual	Mean Error	Percent Error	Residual	Mean Error	Percent Error
Normal1	0.13	-0.04	1.91	0.12	-0.04	1.93	0.07	-0.02	0.93
Normal2	0.08	-0.02	0.94	0.07	-0.03	1.46	0.08	-0.01	0.47
OA1	0.09	-0.04	1.98	0.11	-0.07	3.52	0.08	0.05	2.51
OA2	0.07	-0.02	1.01	0.15	-0.08	4.06	0.12	-0.02	1.02
OA3	0.07	0.05	2.48	0.11	-0.07	3.59	0.08	0.05	2.53
OA4	0.05	-0.1	5.15	0.16	-0.12	6.22	0.11	-0.02	1.06

TABLE 3: Accuracy of cartilage thickness measurement.

Table 4 shows the overall accuracy in cartilage volume measurements with RMS residual difference and RMS mean error.

Clinical symptom	Accuracy of cartilage volume measurement in mm^3								
	Lateral			Medial			Patella		
	Residual	Mean Error	Percent Error	Residual	Mean Error	Percent Error	Residual	Mean Error	Percent Error
Normal1	12.17	-9.40	0.45	8.16	1.10	0.05	9.69	8.90	0.76
Normal2	11.31	5.20	0.24	9.42	-3.40	0.16	8.45	6.70	0.58
OA1	5.65	8.50	0.42	8.71	-1.70	0.08	5.72	2.70	0.23
OA2	7.06	-6.40	0.32	8.64	-6.10	0.30	10.47	9.90	0.88
OA3	5.28	-3.20	0.16	11.34	-6.50	0.33	9.91	-0.20	0.02
OA4	7.47	6.30	0.33	8.65	-7.70	0.40	5.99	5.10	0.52

TABLE 4: Accuracy of cartilage volume measurement.

The thickness measurements of radial search method are compared with standard measurements obtained from edge based method for all the subjects of the data set. The correlation between the cartilage thickness readings of these two methods for entire data set is in good agreement with correlation coefficient varying in the range of 0.91 to 0.97. The cartilage

thickness measurements of these two methods are plotted in Fig. 4(a) for a knee joint and computed correlation is 0.97. This indicates the good accuracy of radial search method. The deviation of cartilage thickness measurements with standard method is found less for a knee joint as indicated by the scattered plot in Fig. 4(b).

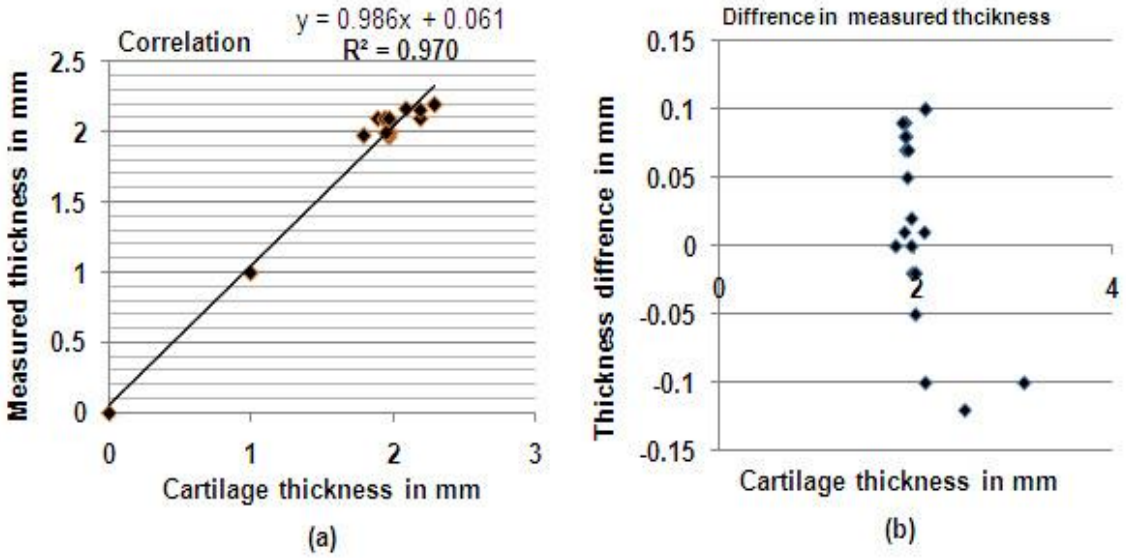


FIGURE 4: Comparison of thickness quantification results.

The thickness measured in a cartilage region for all 7 normal subjects are averaged. The cartilage thickness measured in a cartilage region for all 48 OA affected subjects are averaged. The average cartilage thickness of normal and OA affected are shown in bar chart of Fig. 5(a). The region wise cartilage volumes of normal and OA affected subjects is computed and shown in Fig. 5(b) for comparison.

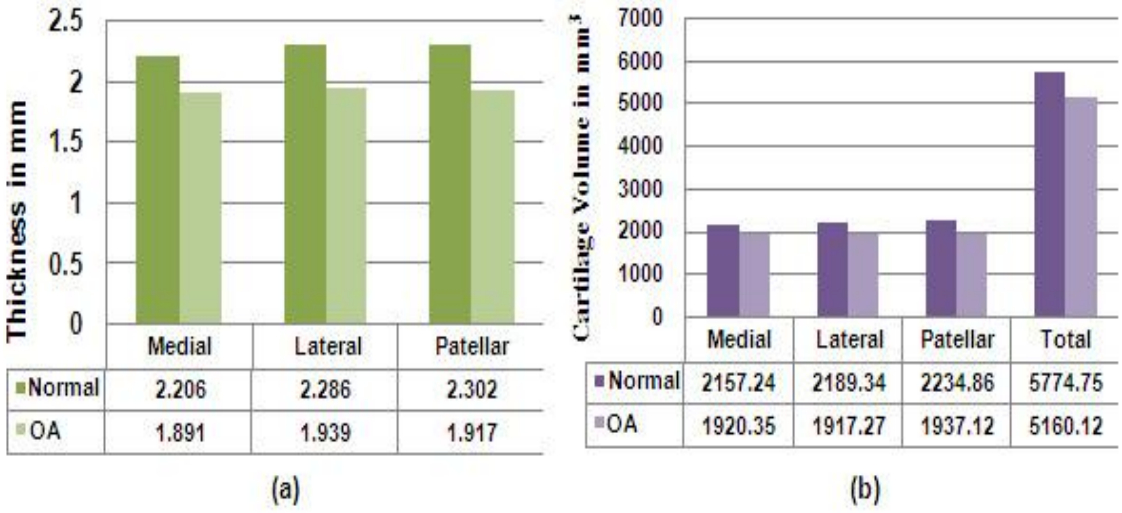


FIGURE 5: Cartilage quantification (a) average thickness (b) average volume.

The cartilage is reconstructed in 3D using texture mapping technique. The different view of visualized cartilage in 3D is shown along with its 2D visualization in Fig. 6.

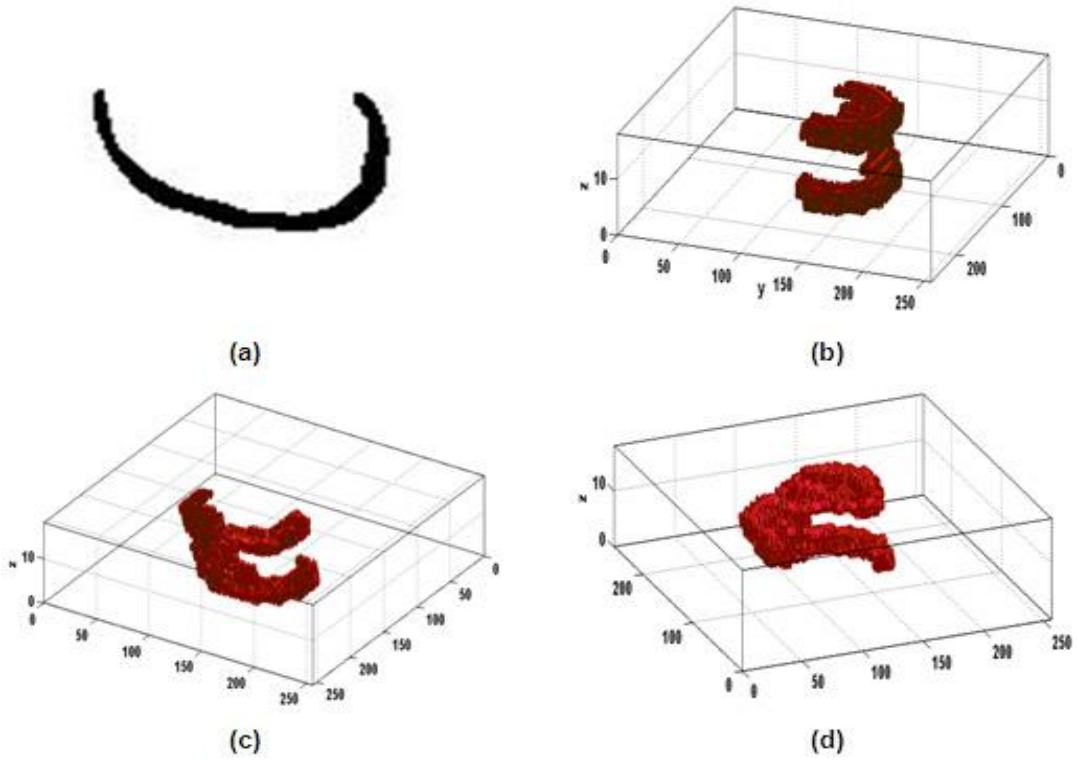


FIGURE 6: Cartilage visualization (a) cartilage in 2D (b)-(d) different views of cartilage in 3D.

4. DISCUSSION

A semiautomatic segmentation method is developed for knee cartilage segmentation from MR images of knee joint. The segmentation method is based on radial search algorithm. The method obtains the boundary points of cartilage by searching along the radial lines from marked origin. There is a modification in search procedure of radial search algorithm compared to earlier reported work [10]. In radial search algorithm the search for the cartilage interfacing boundary starts from the selected origin. In this work, the search starts *r* number of pixels away from the origin along radial line (Fig.2). The incremental angle for search lines is 5°; therefore 60 boundary points are possible for the search range of 0 to 300°. The search starts after 50 pixels along the radial line from the origin; therefore it saves more than 300 pixels to be searched for boundary points in each image of MRI sequence. During the search of outer boundary, the search starts 150 pixels away along each radial line towards the origin. This reduces the number of pixels searched for the detection of outer boundary. The number of samples on the boundary is further increased using B-spline interpolation. Selecting the origin of radial search is the only manual step in segmentation of cartilage. The total time required for processing all the images of 3D MRI sequence takes less than 5 minutes. The processing time of this method is considerably less in comparison with other semiautomatic methods.

The thickness of cartilage is measured in lateral, medial and patellar regions. The arithmetic mean of cartilage thickness is calculated in individual MR images processed. The average of cartilage thickness and SD in a region is calculated as precision measure. Out of the 55 subjects

of this study, results of few normal and OA cases are shown (Table 1). The SD is less than 0.24 in all the cases. This indicates good precision of the readings of cartilage thickness. The COV is less than 5 in normal cases and increases from 3.47 to 10.97 as the level of degradation of cartilage increases in OA cases. The COV indicates the severity of the cartilage degradation. The volume is computed in different compartments of cartilage. Initially, local volumes of cartilage are measured in consecutive images of MRI sequence region wise. To measure the repeatability of the reading, the procedure is repeated thrice on each image. The mean of the cartilage volume from these readings is calculated along with SD (Table 2). The SD is more in OA cases compared to normal cases because of the fact that thickness variations in OA are found to be abrupt. Total volume of entire femur cartilage is computed as summation of volumes in different compartments with ±SD.

The MR images of the dataset are also processed using another validated semiautomatic method of segmentation based on manual marking of boundary, edge detection, and B-spline interpolation [20]. The measurements of this method are taken as standard thickness and volume. The overall accuracy of the measurements is determined as RMS residual difference and RMS error. The RMS residual difference is less than 0.16 mm and error is less than 6.25% for cartilage thickness (Table 3). The RMS residual difference is less than 12.18mm^3 for cartilage volume and error is less than 0.89% (Table 4).

The cartilage thickness measurements obtained using radial search are compared with readings of edge detection based method as standard. The comparison shows good correlation (0.91 to 0.97) between these two methods (Fig.4a), and deviation of cartilage thickness measurements from actual value is found to be very less (Fig.4b). All the 55 subjects of data set are processed for cartilage thickness and volume measurement. The average cartilage thickness and volume of 7 normal and 48 OA cases is calculated for comparison (Fig. 5a and Fig.5b). The result indicates decreased thickness and volume in OA cases compared to normal cases. The segmented cartilage images from MRI sequence are saved and volume rendered for visualization in 3D (Fig. 6). The 3D cartilage is rotated in different directions to visualize different regions.

The modified radial search algorithm used to segment cartilage takes reduced processing time with reduced search area. The quantification of thickness and volume in normal and OA cases is carried out with precision indication. The measurements are compared with another semiautomatic method and the results are found to be in good agreement ($R^2 > 0.91$). The segmented cartilage from knee joint MRI is visualized in 2D and 3D. The reduced thickness and volumes are observed in OA affected knee joints compared to normal knee joints. The computed thickness and volume of articular cartilage is useful in the diagnosis and progressive study of OA affected patients. The results obtained agree with prior clinical diagnosis data.

The menisci (semi-lunar cartilages) at knee joint act as shock absorbers and distribute the forces of weight on the joint surfaces and provide the greater joint stability. Image processing algorithms based on MR images are useful in segmentation of menisci from knee joint. Visualization and quantification of menisci thickness, volume and detection of tears along with cartilage information is useful in diagnosis and treatment of knee abnormalities.

ACKNOWLEDGMENTS
JSS Medical College and Hospital, Mysore, Karnataka, India for providing knee MR Images

5. REFERENCES

[1] Sharma M.K., Swami H.M., Bhatia V., Verma A., Bhatia S.P.S. and Kaur G., “An epidemiological study of correlates of osteoarthritis in geriatric population of UT Chandigarh”, *Indian Journal of Community Medicine*, vol. 32, pp.77-8, 2007.

[2] Reva C.L., David T.F., Charles G.H., Lesley M.A., Hyon Choi, Richard A.D., Sherine Gabriel, Rosemarie Hirsch, Marc C.H., Gene G.H., Joanne M.J., Jeffrey N.K., Hilal Maradit K. and Frederick Wolfe, "Estimates of the prevalence of arthritis and other rheumatic conditions in the United States", *Arthritis & Rheumatism,* vol.58, pp. 26–35, 2008.

[3] Lawrence R.C., Helmick C.G., Arnett F.C., Deyo R.A., Felson David T., Giannini E.H., Heyse S.P., Hirsch R., Hochberg Marc C., Hunder G.G., Liang M.H., Pillemer S.R., Steen V.D. and Wolfe F., "Estimates of the prevalence of arthritis and selected musculoskeletal disorders in the United States", *Arthritis & Rheumatism*, vol. 41(5), pp. 778-799, 1998.

[4] Hussain Z.T. and Usha S. Sinha, "Automated image processing and analysis of cartilage MRI: enabling technology for data mining applied to osteoarthritis", *Proc. Conf. American Institute of Physics,* vol.953, 2007, pp. 262-276.

[5] Peter R.K., Johan L.B., Ruth Y.T.C., Naghmeh R., Frits R.R., Rob G.N., Wayne O.C., Marie-Pierre Hellio Le G. and Margreet K. "Osteoarthritis of the knee: association between clinical features and MR imaging findings", *Radiology,* vol.239, pp. 811-817, 2006.

[6] Stefan M., Tallal C.M., György V., Christoph R. and Siegfried T., "Magnetic resonance imaging for diagnosis and assessment of cartilage defect repairs", *Injury, Int. J. Care of the Injured, v*ol.39S1, pp.S13–S25, 2008.

[7] Thomas M.L., Lynne S.S., Srinka G., Michael R., Ying Lu, Nancy L. and Sharmila M., "Osteoarthritis: MR Imaging findings in different stages of disease and correlation with clinical findings", *Radiology*, vol. 226, pp. 373–381, 2003.

[8] Snoeckx A., Vanhoenacker F. M., Gielen J. L., Van Dyck P. and Parizel P. M., "Magnetic resonance imaging of variants of the knee", *Singapore Med Journal*, vol. 49(9), pp. 734-744, 2008.

[9] Peter M.M. Cashman., Richard I. Kitney, Munir A.G. and Mary E.C., "Automated techniques for visualization and mapping of articular cartilage in MR images of the osteoarthritic knee: a base technique for the assessment of microdamage and submicro damage", *IEEE Trans. Nanobioscience,* vol. 1, pp. 42-51, 2002.

[10] Zohara A. Cohen, Denise M.M., S. Daniel Kwak., Perrine L., Fabian F., Edward J.C., and Gerard A.A., "Knee cartilage topography, thickness, and contact areas from MRI: in-vitro calibration and in-vivo measurements", *Osteoarthritis and Cartilage,* vol.7, pp. 95–109, 1999.

[11] Poh C.L. and Richard I.K., "Viewing interfaces for segmentation and measurement results", Proc*. of 27th Annual Conf. IEEE Engineering in Medicine and Biology*, Shanghai, China, 2005, pp. 5132-5135.

[12] Julio Carballido Gamio, Jan S. Bauer, Keh-Yang Lee, Stefanie Krause and Sharmila Majumdar, "Combined image processing techniques for characterization of MRI cartilage of the knee", Proc. 27th Annual Conf. IEEE Engineering in Medicine and Biology, Shanghai, China, 2005, pp.3043-3046.

[13] Hackjoon Shim, Samuel Chang, Cheng Tao, Jin-Hong Wang, C. Kent Kwoh and Kyongtae T. Bae, "Knee cartilage: efficient and reproducible segmentation on high spatial-resolution MR images with the semi automated graph-cut algorithm method", *Radiology*, vol. 251, pp. 548-556, 2009.

[14] Jenny F., Erik B.D., Ole F.O., Paola C.P. and Claus C., "Segmenting articular cartilage automatically using a voxel classification approach", *IEEE Trans. Medical Imaging*, vol. 26, pp.106-115, 2007.

[15] Claude K., Pierre G., Benoît G., Alain G., Gilles B., Jean P.R., Johanne M.P., Jean Pierre Raynauld, Johanne M.P., Jean Pierre P. and Jacques A. de G., "Computer aided method for quantification of cartilage thickness and volume changes using MRI: validation study using a synthetic model", *IEEE Trans. Biomedical Engineering,* vol. 50, pp. 978-988, 2003.

[16] Jinshan Tang, Steven Millington, Scott T. Acton, Jeff Crandall, and Shepard Hurwitz, "Surface extraction and thickness measurement of the articular cartilage from MR images using directional gradient vector flow snakes", *IEEE Trans. Biomedical Engineering*, vol. 53, pp.896-907, 2006.

[17] Yin Yin, Xiangmin Zhang, Rachel Williams, Xiaodong Wu, Donald D. Anderson and Milan Sonka, "LOGISMOS-Layered Optimal Graph Image Segmentation of Multiple Objects and Surfaces: cartilage segmentation in the Knee Joint", *IEEE Trans. Medical Imaging*, vol. 29, pp. 2023-2037, 2010.

[18] Pierre Dodin, Jean Pierre Pelletier, Johanne Martel Pelletier and François Abram, "Automatic human knee cartilage segmentation from 3D magnetic resonance images", *IEEE Trans. Biomedical Engineering*, vol. 57, pp. 2699-2711, 2010.

[19] Jose G. Tamez Pena, Joshua Farber, Patricia C. Gonzalez, Edward Schreyer, Erika Schneider, and Saara Totterman, "Unsupervised segmentation and quantification of anatomical knee features: Data from the Osteoarthritis Initiative", *IEEE Trans. Biomedical Engineering*, vol. 59, pp.1177-1186, 2012

[20] Mallikarjunaswamy M. S. and Mallikarjun S. Holi, "Segmentation, visualization and quantification of knee joint articular cartilage using MR images", *Springer Multimedia Processing, Communication and Computing Applications*, Proc. first Int. Conf. ICMCCA, 13-15 Dec 2012, vol.213, pp.TP15/1-12.

Comparative Structural Analysis of Phospholipase A2 and Combinatorial Screening of PLA2 Inhibitors

Sanjay Sharma Timilsina , *sanjaytimilsina@gmail.com*
Dept. of Biotechnology
The Oxford College of Science,
H. S. R. Layout,
Bangalore -560 102,
Karnataka, India.

Sarnim Gurung *nimzdude23@gmail.com*
Dept. of Biotechnology,
The Oxford College of Science,
H. S. R. Layout,
Bangalore -560 102,
Karnataka, India.

Roshan Adhikari *roshan028004@gmail.com*
Dept. of Biotechnology
The Oxford College of Science,
H. S. R. Layout,
Bangalore -560 102,
Karnataka, India.

Jignesh Savani *jigssavani@gmail.com*
Dept. of Biotechnology,
The Oxford College of Science,
H. S. R. Layout,
Bangalore -560 102,
Karnataka, India.

Mayank Agrawal *agrawal.mayank08@gmail.com*
Dept. of Biotechnology,
The Oxford College of Science,
H. S. R. Layout,
Bangalore -560 102,
Karnataka, India.

Vedamurthy A.B. *vedamurthy15@gmail.com*
Dept. of Biotechnology,
The Oxford College of Science,
H. S. R. Layout,
Bangalore -560 102,
Karnataka, India.

Joy Hoskeri H. *joybioinfo@gmail.com*
Dept. of Bioinformatics,
Kuvempu University,
Jnana Sahyadri,
Shankaraghatta -577 451,
Shimoga, Karnataka, India.

Abstract

Phospholipases A2 (PLA2) enzyme release fatty acids from the second carbon group of glycerol. This particular phospholipase specifically recognizes the Sn-2 acyl bond of phospholipids and catalytically hydrolyzes the bond releasing arachidonic acid and lysophospholipids. PLA2 are commonly found in mammalian tissues as well as in insects and snakes venom. Venoms constitute a rich source of phospholipase A2 (PLA2) enzymes, which show remarkable diversity in their structure and function. In this investigation, we have made an attempt in analyzing the identical active domain in different PLA2 protein structure isolated from different venoms by studying the conserved active pocket residues. The 21 crystal structures of different PLA2 enzymes isolated from venoms of different species were studied and collected from PDB database. Comparative studies to analyse the conserved active site in this protein was carried out by superimposition studies using TOPMATCH server. To validate the superimposition results sequence alignment studies was carried out using T-COFFEE by multiple sequence alignment analysis. This revealed that 9 PLA2 enzymes from different venoms viz., *Daboia russellii, Cerrophidion godmani, Dienagkistrodon acutus, Bothrops Neuwied, Agkistrodon contortrix, Naja sagittifera, Bos Taurus, Notechis sentatusscutatus* and *Apis mellifera* showed similarity in their active pocket residues, indicating a single drug can effectively occupy their pocket and inhibit the functions of these nine proteins. Hence, in-silico drug designing studies for antivenom drugs against PLA2 was carried out by combinatorial screening of 18 antivenom compounds by docking with PLA2 molecule using Autodock 3.0 tool. In-silico drug designing studies revealed that among 18 antivenom compounds, Indole was most potent in its action in inhibiting the PLA2 function with inhibition constant of 0.04.

Keywords: Phospholipase A2, Antivenom Drugs, Superimposition Studies, Sequence Alignment, Combinatorial Screening, Molecular Docking.

1. INTRODUCTION

Phospholipases A2 (PLA2s) represents an important class of heat-stable, calcium-dependent enzymes catalyzing the hydrolysis of the 2-acyl bond of 3-n-phosphoglycerides. PLA2 enzyme releases the fatty acid from the second carbon group of glycerol. This phospholipase enzyme specifically recognizes the sn-2 acyl bond of phospholipids and catalytically hydrolyzes the bond releasing arachidonic acid and lysophospholipids. Upon downstream modification by cyclooxygenases, arachidonic acid is modified into active compounds called eicosanoids, which are categorized as inflammatory mediators.[1] The Ca^{2+} ion, an essential cofactor, and an Asp residue at position 49 are required for catalysis on artificial substrates. [2] Their catalytic activity upon cell membranes of specific tissues suggests an important role of these enzymes in venoms toxicity. PLA2 are commonly found in mammalian tissues as well as venomous insects, fish venom, frog venom and snake venom.[3] PLA2 exhibits wide varieties of pharmacological effects such as neurotoxicity, cardiotoxicity, myotoxicity, necrosis, anticoagulation, hypotensivity, hemolysis, haemorrhage and edema.[4] Envenomation due to venomous insect or snakes is a serious medical problem, especially in the farms where venomous insects and snakes are abundant. The common species such as *Crotalu rhodostoma, Trimeresurus albolabris, Daboia russeli siamensis, Naja atra* and other venomous insects are responsible for envenomation in Southeast Asia. [5,6] In India itself, on an average 2,50,000 envenomation are recorded in a single year. India is the richest source for poisonous species of frogs, insects and snakes. Among these majority of the bites and mortality are attributed to snake species like *Ophiophagus hannah* (King cobra), *Naja naja* (Spectacled cobra), *Daboia russelli* (viper), *Bungarus caeruleus* (Common krait), *Echis carinatus* (Saw scaled viper) etc. Among them most reports are on haemolytic venomous snakes. These snake haemolytic venoms constitute a rich content of PLA2 enzymes. Snake venom phospholipase A2 (svPLA2) can induce several additional pathophysiological effects such as cardiotoxicity, myotoxicity, pre or postsynaptic neurotoxicity, edema, hemolysis, hypotension, convulsion, platelet aggregation inhibition and

anticoagulation. [7-11] Venoms from insects, frogs, snakes, fishes constitute a rich source of PLA2 enzymes, which show remarkable functional similarity. Although, PLA2 enzyme is found in various venom derived from different organism, but possesses similarity in their substrate specificity. This reveals a hypothetical basis that this similarity is due to the conserved active site in these enzymes. In continuation with our interest to unmask the molecular basis of structural similarity and functional identity, we carried out an bioinformatics approach to study 21 PLA2 structures derived from various organisms viz*., Daboia russellii, Bungarus caervleus, Sos scrofa, Bothrops jararacussu, Naja sagittifera, Bos Taurus, Bothrops neuwiedi, Naja sagittifera, Bothrops neuweidi pauloensis, Vipera ammodytes meridionalis, Micropechis ikaheka, Agkistrodon contortrix, Echis carinatus, Ophiophagus hannah, Streptomyces violaceoruber, Gloydius halys, Dienagkistrodon acutus, Notechis sentatus scutatus, Apis mellifera, Cerrophidion godmani, Crotalus atrox* by structure superimposition studies to understand and confirm the structural relationship and analyse the conserved active site. Further, rigorous literature survey revealed that there are no reports available on the active pocket residues of PLA2 enzyme. Hence, we efficiently elucidated the active site and the residues that fall within this domain by bioinformatics approach. Combinatorial screening of 18 antivenom compounds was carried out by docking them in the elucidated active pocket by molecular docking approach to shortlist the potent compounds that can act as an effective inhibitor against all the PLA2 enzyme that possess similar active pocket.

2. MATERIALS AND METHODS

2.1. Collection of PLA2 Crystal Structures

The crystal structures of PLA2 enzyme isolated from venoms of different species were studied and collected from Protein Data Bank (PDB) database. Out of different structures present in the PDB database (www.rcsb.org/pdb), one PLA2 structure of each species present was downloaded in pdb format. PLA2 structure of *Daboia russellii* (with PDB ID - 3CBI) as a parent molecule was downloaded along with its sequence. Similarly PLA2 structures with PDB ID 2OSN (*Bungarus caervleus*), 2AZY (*Sos scrofa*) , 1ZL7 (*Bothrops jararacussu*), 1TD7 (*Naja sagittifera*), 2BAX (*Bos Taurus*), 1PCQ (*Bothrops neuwiedi*), 1ZM6 (*Naja sagittifera*), 1PC9 (*Bothrops neuweidi pauloensis*), 1Q5T (*Vipera ammodytes meridionalis*), 1OZY (*Micropechis ikaheka*), 1S8G (*Agkistrodon contortrix*), 1OZ6 (*Echis carinatus*), 1GP7 (*ophiophagus hannah*), 1IT4 (*Streptomyces violaceoruber),* 1M8R *(Gloydius halys),* 1MG6 *(Dienagkistrodon acutus),* 2NOT *(Notechis sentatus scutatus),* 1POC *(Apis mellifera),* 1GOD *(Cerrophidion godmani),* 1PP2 *(Crotalus atrox*) were downloaded along with their sequence.

2.2. Structure Superimposition Studies

Structural comparison by superimposition studies was carried out by using TOPMATCH server to elucidate the conserved active site in PLA2 enzyme isolated from different organisms. [12] In this approach the PLA2 structure of *Daboia russellii* (with PDB ID - 3CBI) was used as the parent protein molecule (target) and other PLA2 structures derived from other organisms were used as query and superimposed over this parent protein molecule. The superimposed PLA2 target and query protein structures were viewed using Jmol viewer online (www.jmol.org). After the TOPMATCH was performed the superimposed structure were downloaded in PDB format and the sequence alignment were downloaded in PDF format.

2.3. Multiple Sequence Alignment Studies

PLA2 protein sequences of *D. russellii, B. caervleus, S. scrofa, B. jararacussu, N. sagittifera, B. taurus, B. neuwiedi, N. sagittifera, B. neuweidi pauloensis, V. ammodytes meridionalis, M. ikaheka, A. contortrix, E. carinatus, O. hannah, S. violaceoruber, G. halys, D. acutus, N. sentatus scutatus, A. mellifera, C. godmani* and *C. atrox* were collected in FASTA format from Protein Data Bank. PLA2 protein sequences similarity and availability of conserved signatures in all the 21 protein sequences was interpreted and studied by multiple sequence alignment approach using T-Coffee server (www.ebi.ac.uk/tools/msa/tcoffee) [13]. Multiple sequence alignment (MSA) file was downloaded and viewed in a color format using JalView. Results obtained by MSA study was

used to support the superimposition study and to identify active domain and the amino acid residues that fall within this domain.

2.4. PLA2 Enzyme Active Site Elucidation and Identification of Active Site Residues

The active pocket of PLA2 protein molecule was elucidated by using PDBSum of RCSB server (www.ebi.ac.uk/pdbsum/). In this approach, 17 PLA2 structures along with their inhibitors were randomly selected and studied for the amino acid residues which were interacting with the inhibitor compound as revealed by PDBSum protein-ligand interaction map. The selected inhibitor molecules varied in their interaction position. The amino acid residues that were commonly involved in interaction with all the inhibitor molecules were considered to be conserved and were representing the active site that is found to be occupied by almost all the inhibitors. This hypothesis was emphasized in elucidating the active pocket. Taking into account the number of common amino acid residues of these 17 PLA2 interacting with inhibitor ligands, the active pocket of PLA2 molecule was determined and these common interacting residues were considered as active site residues.

2.5. Study and Collection of Antivenom Drugs

Rigorous literature survey was carried out to short list potent antivenom and PLA2 inhibitors. Altogether 18 different antivenoms/PLA2 inhibitors viz., Acenapthene, aspirin, atropine, betasitosterol, diisobutyl phthalate, diosgenin, indole, isooxazolidine, monalide, petrosaspongiolide, piperazine, quercitin, revesatrol, rosmarinic acid, rutin, tetracycline, vidadol and wedelolactone were selected and their structures were downloaded. The 3D coordinate's file of these 18 ligand molecules was generated using Prodrg (davapc1.bioch.dundee.ac.uk/prodrg/).[14] The 3D coordinate file compatible for Autodock 3.0 was generated and downloaded in (.pdbq) format. The 3D coordinate files of all 18 molecules were used for combinatorial screening.

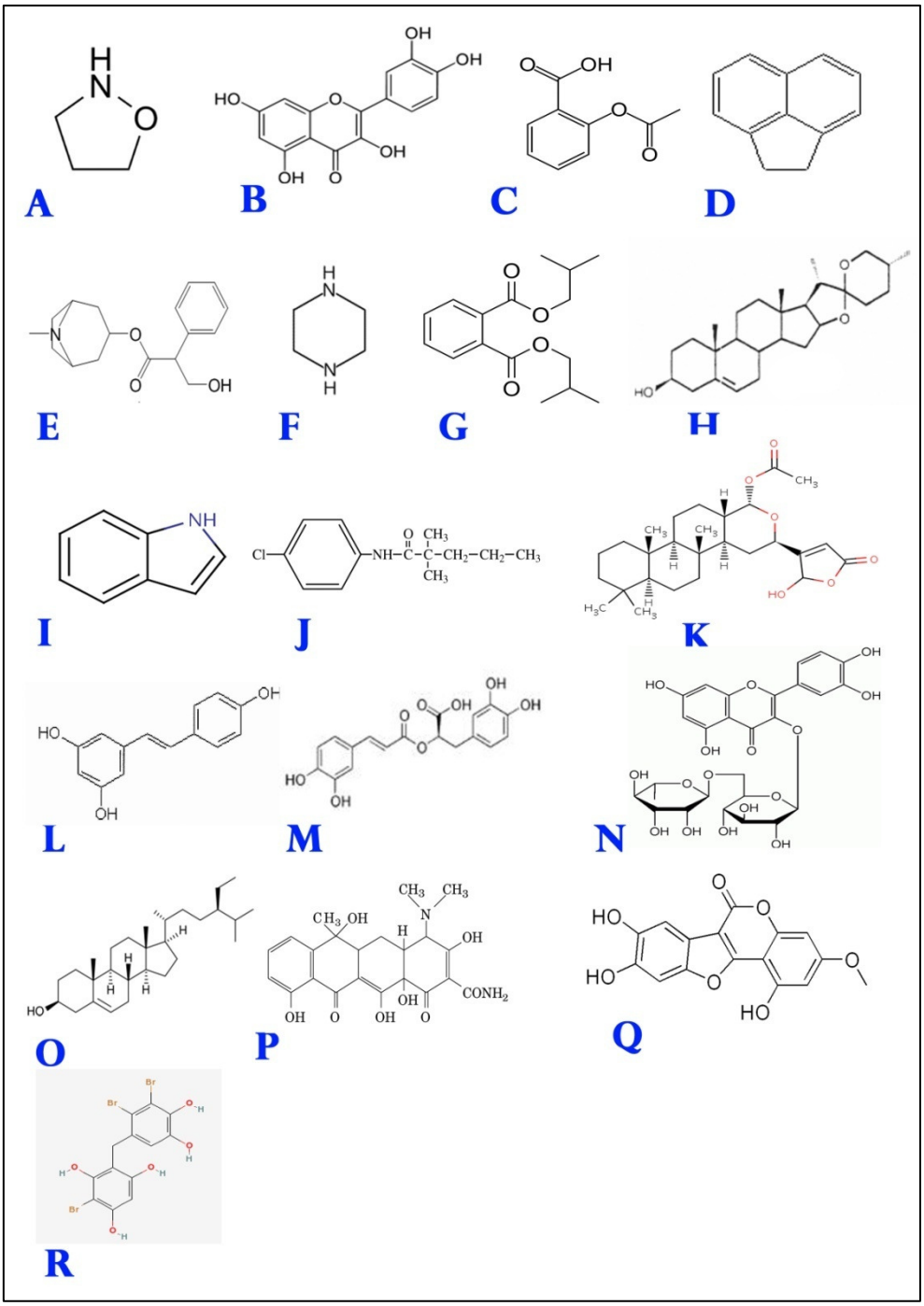


FIGURE 1: Short listed potent antivenom and PLA2 inhibitors viz., (A) Isoxazolidine, (B) Quercetin, (C) Aspirin, (D) Acenaphthalene, (E) Atropine, (F) Piperazine, (G) Di-isobutylphthalate, (H) Disogenin, (I) Indole, (J) Monalide, (K) Petrosaspongiolide, (L) Revesatrol, (M) Rosamarinic acid, (N) Rutin, (O) Betasitosterol, (P) Tetracyclene, (Q) Wedelolactone, (R) Vidalol for combinatorial screening.

2.6. Combinatorial Screening by Molecular Docking Study

Automated docking was used to carry out combinatorial screening and determine the orientation of all 18 antivenom compounds bound in the active site of PLA2 protein from *Daboia russellii*. A lamarckian genetic algorithm implemented in the AutoDock 3.0 program was employed. [15] All the 18 antivenom compounds were designed and the structure was analyzed by using ChemDraw Ultra 6.0. 3D coordinates were prepared using PRODRG server. [16] The protein structure file 3CBI was taken from PDB (www.rcsb.org/pdb), was edited by removing the heteroatoms and C terminal oxygen was added by using SPDBV tool. [17] For docking calculations, GasteigereMarsili partial charges [18] were assigned to the ligands and nonpolar hydrogen atoms were merged. All torsions were allowed to rotate during docking. The active pocket residues were predicted using CASTp server. [19] The grid map, which was centered at the following active pocket residues of the protein 3CBI [LEU2(A), GLY30(A), HIS48(A), ILE19(A), TRP31(A), ASP99(A), LYS69(A), TYR52(A), SER23(A), TYR22(A), ASP49(A), PHE5(A), ALA18(A)], and generated by applying AutoGrid. The lamarckian genetic algorithm was applied for minimization, using default parameters. The number of docking runs was 10, the population in the genetic algorithm was 250, the number of energy evaluations was 100,000, and the maximum number of iterations 10,000.

3. RESULTS AND DISCUSSION

PLA2 is commonly found in mammalian tissues, venomous insects, fish venom, frog venom and snake venom but PLA2 from different organisms possesses similarity in their substrate specificity. Based on this knowledge, the similarity in their function and substrate specificity is due to the conserved active domain. This study was undertaken to analyze and unmask the molecular basis of structural similarity and functional identity. In the present investigation, we carried out a bioinformatics approach to study 21 PLA2 structures derived from various organisms by structure superimposition studies to understand and confirm the structural relationship and analysis the conserved active site to elucidate the active pocket residues that fall within this active domain. Further combinatorial screening of 18 antivenom compounds was carried out by docking the ligands in the elucidated active pocket by molecular docking approach.

3.1. Multiple Sequence Alignment Studies

Multiple sequence alignment by t-coffee revealed that protein with PDB ID 1IT4 (violaceoruber), 1PCQ (Bothrops neuwiedi) and 1POC (Apis mellifera) did not show any relationship with the other protein sequence in the dataset. Whereas, other protein molecule showed region specific similarity. The region segment that showed maximum identity in the data set is depicted in Figure 2.

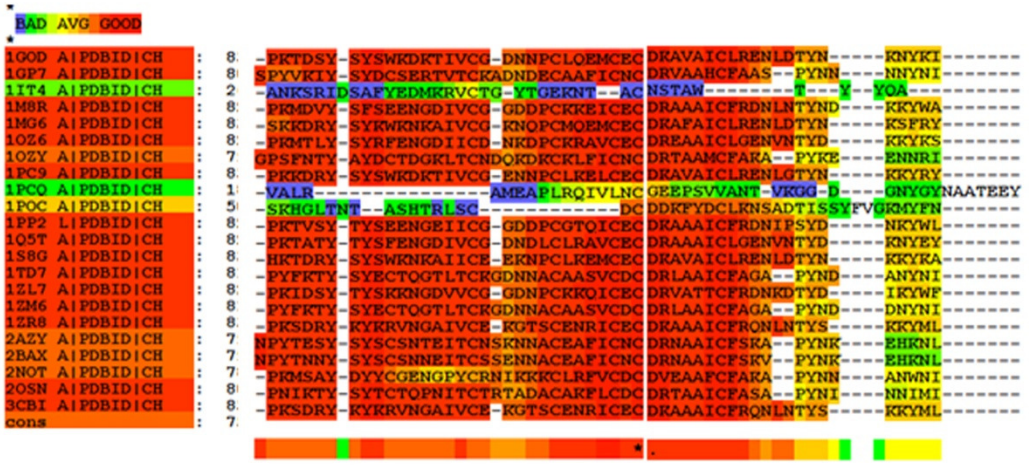


FIGURE 2: Multiple sequence alignment showing the conserved region among the 21 PLA2 protein molecules under study.

3.2. Structure Superimposition Studies

Structure superimposition study using TOPMATCH server was carried out for 21 different PLA2 structures with 3CBI of *Daboia russelli* as a target reference molecule, to which all other PLA2 structures were superimposed to identify and interpret the structurally identical region among the PLA2 enzymes. After a single run of superimposition of each PLA2 enzyme structure against the parent target PLA2 with PDB ID 3CBI, the superimposed structures were downloaded in pdf format and interpreted. Results of structure superimposition are depicted in figure 4. Further, sequence superimposition in support to structural superimposition also demonstrated similar results. The result of structural superimposition and sequence superimposition study revealed that among all the PLA2 structures used for the study, the PLA2 structures with PDB ID as 1GOD, 1MG6, 1PC9, 1S8G, 1ZM6, 2BAX, 2NOT, 1POC showed high degree of similarity in the first 122 amino acid in chain A, this is the frame within which the active pocket residue are found and hence this frame represents the active site of PLA2 (Figure 3). Further, this investigation hypothesizes the structural similarity within the active site domain reveals their functional similarity and substrate specificity. However, four PLA2 molecule with PDB ID: 2OSN, 2AZY, 1QST, 1OZ6 also showed superimposition match in last 122 amino acid sequence, but this frame is out of the active site domain.

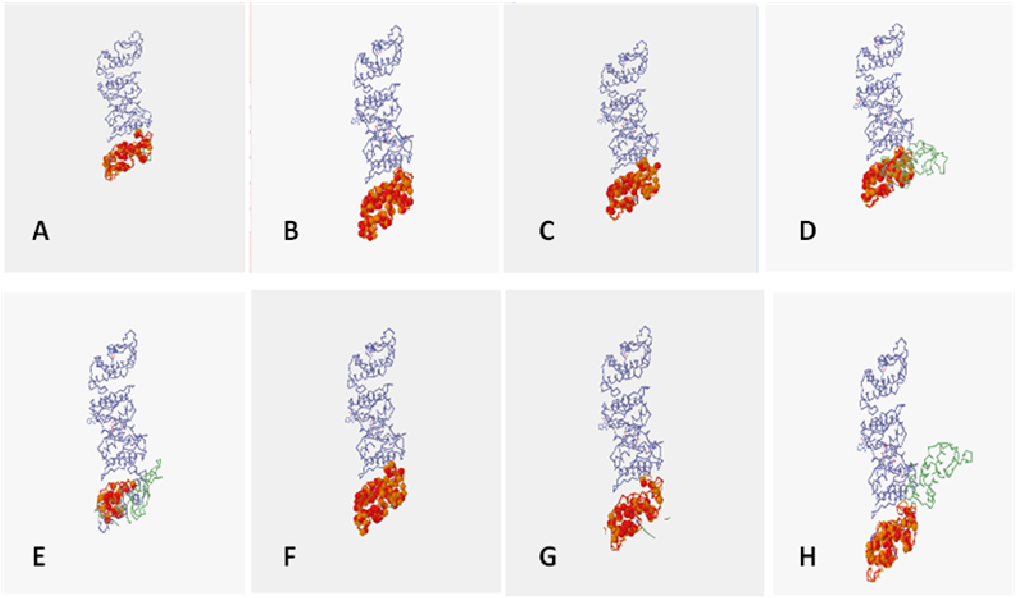


FIGURE 3: Superimposed structures showing superimposed region of 3CBI (Target) with query (A) 1GOD, (B) 1MG6, (C) 1PC9, (D) 1S8G, (E) 1ZM6, (F) 2BAX, (G) 2NOT and (H) 1POC.

This investigation support the hypothesis that, PLA2 enzyme although exists in various organism, either in venom or in mammalian tissue, their functional similarity is always documented. Through this investigation, it is presented that 8 PLA2 enzymes from organism viz., *Cerrophidion godmani, Dienagkistrodon acutus, Bothrops neuweidi pauloensis, Agkistrodon contortrix, Naja sagittifera, Bos Taurus, Notechis sentatus scutatus* and *Apis mellifera* showed active site domain structural similarity with PLA2 of *Daboia russelli*. This is in direct corroboration with the fact *that* active pocket's structural similarity dictates their functional similarity.

FIGURE 4: Amino acid sequence frame that falls within the superimposed region of 3CBI (Target) with query 1GOD, 1MG6, 1PC9, 1S8G, 1POC, 1ZM6, 2BAX and 2NOT.

3.3. PLA2 enzyme active site elucidation and identification of active site residues

The active pockets of PLA2 molecule of *Daboia russelli* which was used for docking study was identified using PDBSum of RCSB server. 17 different PLA2 molecule of *Daboia russelli* were randomly selected and studied for the ligand interaction. The interacting amino acid residues with different ligands, their repetition in different PLA2 structures and their respective frequency is as noted in the table. Taking in consideration the frequency of their repeat the active pocket of PLA2 of *Daboia russelli* was considered to be the domain containing following residue (Leu2(A), Phe5(A), Ala18(A), Ile19(A), Tyr22(A), Ser23(A), Gly30(A), Trp31(A), His 48(A), Asp49(A), Tyr52(A), Lys69(A), Asp99(A)).

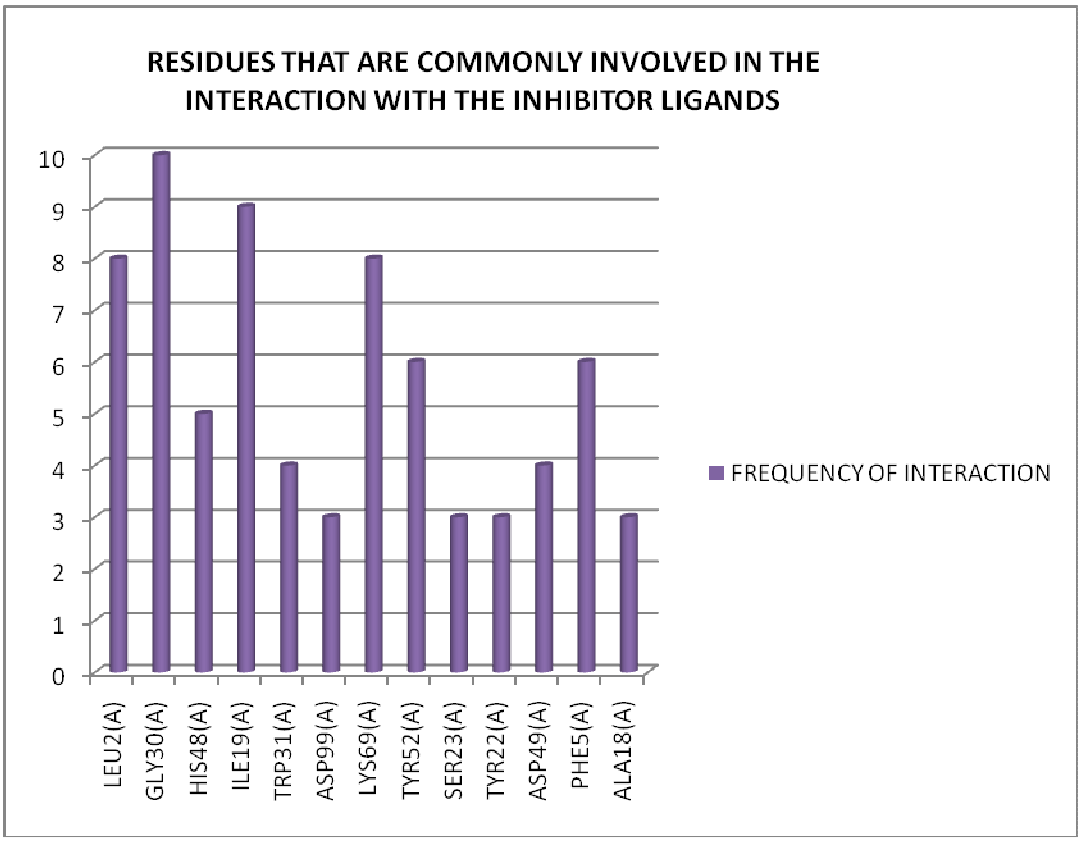


FIGURE 5: Graph showing the frequency of interaction of the residues commonly interacting with the inhibitor ligands.

3.4. Combinatorial Screening by Molecular Docking Study

Comparative computational docking of 18 potent antivenom compounds to the PLA2 revealed that Indole, Isooxazolidine, Piperazine, Diisobutyl phthalate, Tetracycline and Rosamarinic acid showed very high inhibition constant in the decreasing order (0.04, 0.01, 0.000627, 0.00546, 0.000259 and 0.000121 respectively) among all the antivenom compounds.

The docked energy for Indole was -10.55 with an estimated inhibition constant of 0.04 and intermolecular energy -2.09, while the docked energy of Isooxazolidine was -3.62 with an inhibition constant of 0.01 and intermolecular energy of1-3.62. Indole didn't show any hydrogen bonding while Isooxazolidine showed hydrogen bonding with the backbone hydrogen of Gly 30 of PLA2.

Results of combinatorial screening revealed that among all 18 antivenom molecules Indole was found to be the most potent antivenom compound. Hence, Indole, Isooxazolidine, Piperazine, Diisobutyl phthalate, Tetracycline and Rosamarinic acid can be used to inhibit all the PLA2 enzymes from *Cerrophidion godmani, Dienagkistrodon acutus, Bothrops neuweidi pauloensis, Agkistrodon contortrix, Naja sagittifera, Bos Taurus, Notechis sentatus scutatus,* and *Apis mellifera* as they show structural similarity with PLA2 of *Daboia russelli*. Hence, they possess ligand or substrate specificity and hypothetically have the ability to bind common compound. Through this investigation it can be concluded that PLA2 enzyme from the above mentioned organisms can be inhibited using a single potent drug like Indole.

compounds	Best Orient -ation	Binding Energy	Docking Energy	Inhibition Constant	Intermol Energy	Torsional Energy	Internal Energy	RMS	No. of H-bonds	H-Bonding with Amino acid Residue
Piperazine	10th	-4.37	-4.37	0.000627	-4.37	0.0	0.0	0.0	0	-
Tetracycline	4th	-4.89	-5.55	0.000259	-5.51	0.62	0.03	0.0	4	SER23, SER23, SER23, SER23.
Rosmarinic acid	8th	-5.34	-7.86	0.000121	-7.83	2.49	-0.03	0.0	0	-
Diisobutyl phthalate	2nd	-4.45	-6.87	0.000546	-6.94	2.49	0.07	0.0	0	-
Indole	4th	-2.09	-2.09	0.04	-2.09	0.0	0.0	0.0	0	-
Isooxazolidine	2nd	-3.0	-3.62	0.01	-3.62	0.62	0.0	0.0	1	GLY30
Rutin	7th	-5.5	-6.37	9.37e-005	-7.05	1.56	0.68	0.0	1	GLY30
Diosgenin	7th	-9.98	-9.98	4.85e-008	-9.98	0.0	0.0	0.0	0	-
Monalide	5th	-6.1	-6.38	3.38e-005	-7.97	1.87	1.59	0.0	0	-
Petrosaspongiolide	8th	-8.73	-9.9	4e-007	-9.66	0.93	-023	0.0	0	-
Quercitin	6th	-9.25	-9.57	1.65e-007	-9.56	0.31	-0.01	0.0	3	LEU2, TYR22, GLY30
Reveratrol	2nd	-7.21	-8.26	5.21e-006	-8.14	0.93	-0.92	0.0	0	-
Wedelolactone	1st	-8.39	-8.72	7.09e-007	-8.7	0.31	-0.02	0.0	0	-
Vidadol	5th	-6.75	-7.44	1.13e-005	-7.37	0.62	-0.07	0.0	1	ASP49
Aspirin	3rd	-5.75	-6.47	6.08e-005	-6.69	0.93	0.21	0.0	2	GLY30, HIS48
Atropine	7th	-6.08	-7.1	3.47e-005	-7.64	1.56	0.54	0.0	2	SER23, GLY30
Betasitosterol	1st	-7.7	-9.49	2.28e-006	-9.56	1.87	0.07	0.0	0	-
Acenaptherene	5th	-7.48	-7.48	3.29e-006	-7.48	0.0	0.0	0.0	0	-

TABLE 1: Results of in silico molecular docking studies of PLA2 inhibitors.

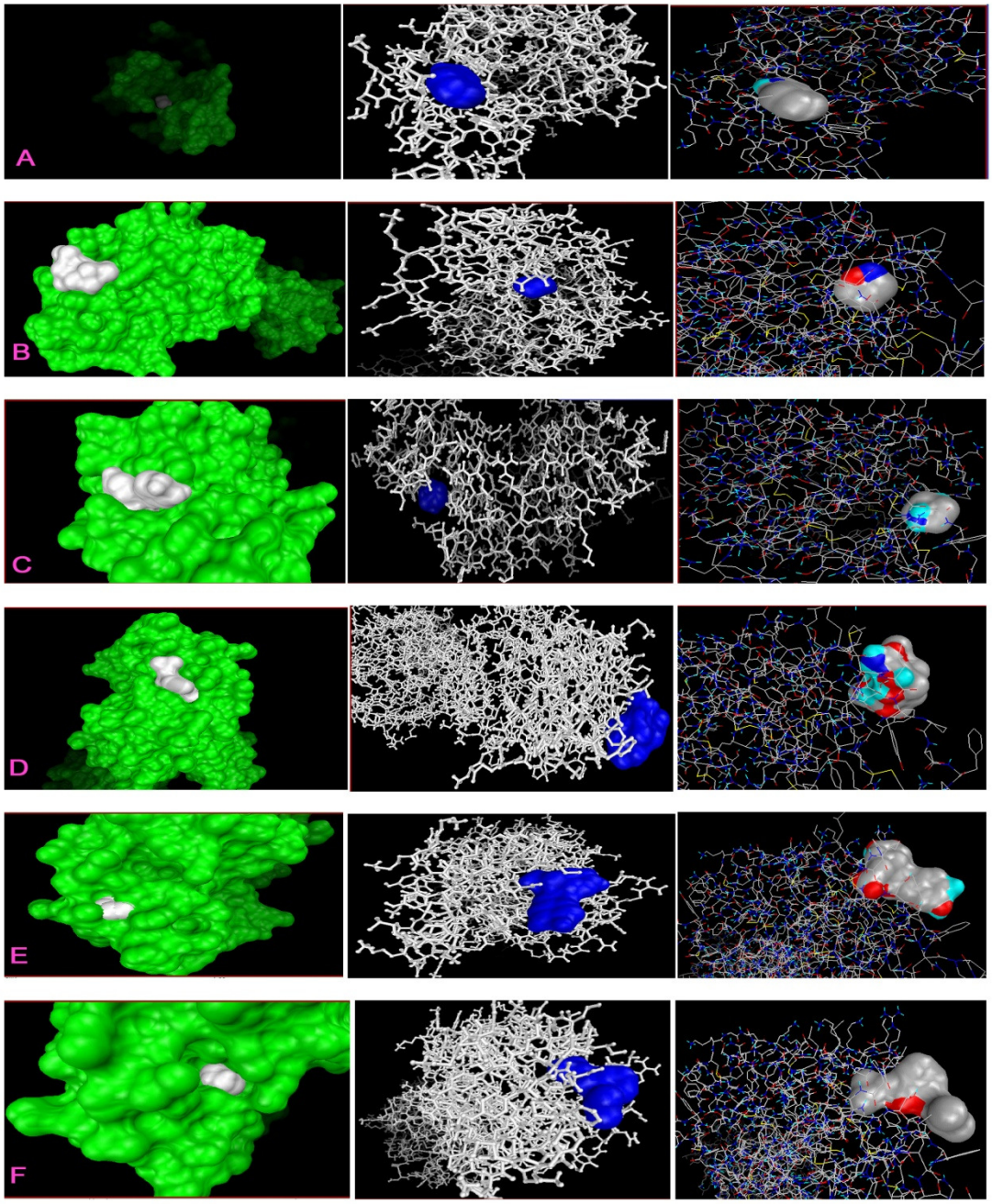


FIGURE 6: Docking of PLA2 with (A) Indole, (B) Isooxazolidine, (C) Piperazine, (D) Tetracycline, (E) Rosmarinic acid, (F) Diisobutyl phthalate.

4. REFERENCES

1. E.A. Dennis. “Diversity of group types, regulation, and function of phospholipaseA2." Journal of Biological Chemistry. Vol 269(18), pp. 13057-13060, May. 1994.

2. Y.H. Pan, T.M. Epstein, M.K. Jain and B.J. Bahnson. "Five coplanar anion binding sites on one face of phospholipase A2: relationship to interface binding". Biochemistry. Vol 40(3), pp. 609–617, Jan. 2001.

3. J.P. Nicolas, Y. Lin, G. Lambeau, F. Ghomashchi, M. Lazdunski and M.H. Gelb. "Localization of structural elements of bee venom phospholipase A2 involved in N-type receptor binding and neurotoxicity". Journal Biological Chemistry. Vol 272(11), pp. 7173-7181, Mar. 1997.

4. R.M. Kini. Venom Phospholipase A2 Enzymes: Structure, Function and Mechanism. England, John Wiley & Sons, Chichester, 1997, pp. 1-511.

5. J. P. Chippaux. Snake bites- appraisal of the global situation. Bulletin of World Health Organization, Vol 76(5), 515-524, May. 1998.

6. L.A. Ferreira., O.B. Henriques, A.A. Andreoni, G.R. Vital, M.M. Campos, G.G. Habermehl, and V.L. de Moraes. “Antivenom and biological effects of ar-turmerone isolated from Curcuma longa (Zingiberaceae)”. Toxicon. 30, pp. 1211–1218, Oct. 1992.

7. A. Argiolas and J.J. Pisano. "Facilitation of phospholipase A2 activity by mastoparans, a new class of mast cell degranulating peptides from wasp venom". Journal of Biological Chemistry. Vol 258(22), pp. 13697-13702, Nov. 1983.

8. Z. Mallat, G. Lambeau, A. Tedgui. "Lipoprotein-Associated and Secreted Phospholipases A2 in Cardiovascular Disease: Roles as Biological Effectors and Biomarkers". Circulation. Vol 122(21), pp. 2183-2200. Nov. 2010.

9. D. De Luca, A. Minucci, P. Cogo, E.D. Capoluongo, G. Conti, D. Pietrini, V.P. Carnielli, and M. Piastra. "Secretory phospholipase A pathway during pediatric acute respiratory distress syndrome: a preliminary study". Pediatric Critical CareMediciine. Vol 12(1), 20-24, Jan. 2011.

10. W.R. Henderson, R.C. Oslund, J.G. Bollinger, X. Ye, Y.T. Tien, J. Xue, and M.H. Gelb. "Blockade of Human Group X Secreted Phospholipase A2 (GX-sPLA2)-induced Airway Inflammation and Hyperresponsiveness in a Mouse Asthma Model by a Selective GX-sPLA2 Inhibitor". Journal of Biological Chemistry, Vol 286(32), 28049–28055, Aug. 2011.

11. Y. Wei, S.P. Epstein, S. Fukuoka, N.P. Birmingham, X.M. Li, and P.A. Asbell. "sPLA2-IIa Amplifies Ocular Surface Inflammation in the Experimental Dry Eye (DE) BALB/c Mouse Model". Investigative Opthalmology and Visual Science. Vol 52(7), pp. 4780–4788, Jul. 2011.

12. M.J. Sippl and M. Wiederstein. “A note on difficult structure alignment problems”. Bioinformatics. Vol 24(3), pp. 426-427, Jan. 2008.

13. C. Notredame, D.G. Higgins and J. Heringa. “T-Coffee: A novel method for fast and accurate multiple sequence alignment”. Journal of Molecular Biology. Vol 302(1), pp. 205-217, Sep. 2000.

14. A.W. Schüttelkopf and D.M.Van Aalten. “PRODRG: a tool for high-throughput crystallography of protein-ligand complexes”. Acta Crystallographica Section D Biological Crystallography. Vol 60(8), pp. 1355-1363, Aug. 2004.

15. R. Bhat, Y. Xue and S. Berg. “Structural insights and biological effects of glycogen synthase kinase 3-specific inhibitor AR-A014418”. Journal of Biological Chemistry. Vol 278(46), pp. 45937–45945, Aug 2003.

16. A.K. Ghose and G.M. Crippen. “Atomic physicochemical parameters for three-dimensional-structure-directed quantitative structure–activity relationships 2. Modeling dispersive and

hydrophobic interactions". Journal of Chemical Information Computer Sciences. Vol 27(1), pp. 21-35, Feb. 1987.

17. T.A. Binkowski, S. Naghibzadeg and J. Liang. " CASTp computed atlas of surface topography of proteins". Nucleic Acid Research. Vol 31(13), pp. 3352–3355, Jul. 2003.

18. J. Gasteiger and M. Marsili. "Iterative partial equalization of orbital electronegativity - a rapid access to atomic charges". Tetrahedron. Vol 36(22), pp. 3219-3288, Mar. 1980.

19. T. Reya and H. Clevers. " Wnt signalling in stem cells and cancer". Nature. Vol 434(7035), pp. 843-850, Apr. 2005.

China Kang Heating Room Fresh Air into the Indoor Research Facilities

Hongguo Ren *renhongguo@126.com*
College of Civil and Architecture Engineering
Northeast Petroleum University; Daqing
Heilongjiang163318; China

Abstract

Based on Chinese fire resistance heating room existing air quality problem, on the basis of investigation and research, detailed arrangement the town heated kang residential air supplement mode, and analyzes the typical example, put forward the improvement fire resistance heating and fresh air into the indoor facilities design scheme, and research into new facilities in rural different forms of heated kang residential application forms.

Key words: Fire Resistance, Heating into the Fresh Air.

1. INTRODUCTION

Fire resistance is the earliest human control method of heating, due to its many advantages has been widely used in cold regions of China. Since the reform and opening up, Chinese rapid development of economic construction, the city of multilayer high-rise residential heating basic use central heating, fire resistance thus gradually from city residential gradually retreat, but in city suburbs and the vast rural areas is still widely used[1]. The country's present energy conservation, the environmental pollution and most of the village economic income perspective, heated kang housing is still irreplaceable ideal heating mode. China's rural house indoor air quality is mainly by cooking, heating and other consumer life energy impacts, villages and towns use clean fuel or senior less fuel, still in biomass fuel and coal is given priority to. The main fuel and cooking, heating stoves use, to indoor send out a CO, SOX, NOX, such as fine particulate air pests[2]. Indoor air pollution (IAP) is a health hazard of the main environmental factor, 4% of the global burden of disease due to indoor air pollution, and tobacco disease burden is quite. In rural areas of indoor air pollution is mainly due to fuel combustion and of generation, especially the incomplete combustion, a large number of pollutants are released into the indoor air, in the winter more serious (twenty times the WHO standard). In China, coal and biological fuel is the main fuel in rural areas, the fuel burns can release particulates (RMP), carbon monoxide, sulfur dioxide, fluoride, aldehydes, etc to the health of human body harmful substances. Developing countries some

monitoring data show that the indoor air fine particulate concentration reaches as high as 1000 to 10000 mg/m3, more than the national environmental protection agency standard (100 mg/m3) 10-100 times. Developing countries rural areas indoor particulate exposure level in rural areas is the developed countries of seven to 26 times. Indoor air pollutants are because respiratory disease (such as acute respiratory infection and chronic obstructive pulmonary disease) of the main risk factors, especially for women and children's health is more serious damage. Around the world each year 12 million children under the age of five die 20% died of acute lower respiratory tract infection. The World Bank report stressed that IAP in China's poor suburb, rural areas and alpine region women and children lung disease burden is very high, the main reason each year about 150000 children under the age of five die in acute lower respiratory tract infection. Reports estimate, in China each year about excess killed 111000 people, 220000 people, in emergency 4.3 million person-time, ' '3 million campaign is limited, due to indoor air pollution. Indoor air pollutants chronic exposure can also lead to low birth weight, increase the incidence of tuberculosis and cataract. Our rural house interior into the fresh air ventilation system research and development, through the questionnaire survey and field testing combination way. The questionnaire according to the particularity of the small towns, in order to emphasize the indoor into the fresh air. Research shows that: the residential indoor environment; Residents to fresh health and energy consumption in understanding; Residential natural into fresh air condition investigation, Natural into fresh air condition, specific include different seasons of different function room window area, open the window way and window time survey; Mechanical into fresh air condition, specific include toilet, kitchen and bedroom indoor ventilation equipment selection condition survey. At the same time, in order to understand the effect of fresh air into the system, but also for indoor air quality and the indicators of the fresh air volume test [3]. Compared with cities, rural house in natural fresh air into a unique advantage. Town of outdoor air quality is better than the city, if can reasonable use of natural into new technology, strengthen interior into fresh air ventilation, small town residents residential indoor environment status will be greatly improved [4].

2. FIRE RESISTANCE HEATING AND INTO THE FRESH AIR FACILITIES STATUS

2.1 Town Heated Kang Residential Air Supplement Mode

Rural house there has always been a winter indoor air quality problems, especially in the national "eleventh five-year plan" project to support science and technology work of the research found that due to the building energy efficiency makes envelop thermal insulation, air performance has been much improved, through the doors and Windows and retaining structure aperture infiltration into building internal outdoor fresh air, lead to reduce indoor air quality deterioration, and building and usually in into the fresh air facilities into consideration is not enough or no consideration, in building energy saving in further today, to pay special attention to residential buildings in cold air into the research, especially in heating season into the problem of indoor air, Otherwise pay a cost to build energy saving building, on the contrary, because indoor air pollution to habitant bring serious harm.

(1) Permeability

Most rural house is breathable, i.e., building shell contains a lot of air inlet and air out of the channel, such as doors and Windows, in and out of the gap, wire around the pipeline gateway, foundation and holes. Outside air into by building insulation shell seal degree, wind speed and the temperature difference between inside and outside and the influence of environmental factors. Due to the building structure and building energy saving technology to promote this way into fresh air is decreasing.

(2) Open the door into the wind

Heated kang housing in the use of time will be DaoYan, shoot, meteor fire (see chart 1), such as the lack of fresh air, leading to the villagers in cooking, burn when kang door are often forced to open, the door into the fresh air is to accelerate the air flow, eliminate indoor pollutants plays a role, but sufficient fuel combustion, human activities on the new requirement is long, and not be able to solve a centralized open, this kind of centralized open but make a lot of heat loss, cause human energy saving failure.

FIGURE 1: ChaiZao Meteor Fire Schematic Diagram.

2.2 Analysis of the Status of Typical Examples

The author research by heilongjiang housing and urban and rural construction and French environment and energy control department design of heihe love bright outside the three DaoGou

village close tree village residential use fire resistance heating and indoor air facilities in residential fire resistance (see chart 2). Based on the case of residential closed consideration, chimney is the only foreign empty. Indoor heated kang, ChaiZao use and human activities will lead to internal pressure is reduced, the outside air chimney reverse pressure into indoor, the heated kang house doesn't burn well, indoor air quality bad based on this thinking puts forward the fire resistance heating indoor into fresh air facilities design patterns.

Both fresh air into the facility design is as follows:

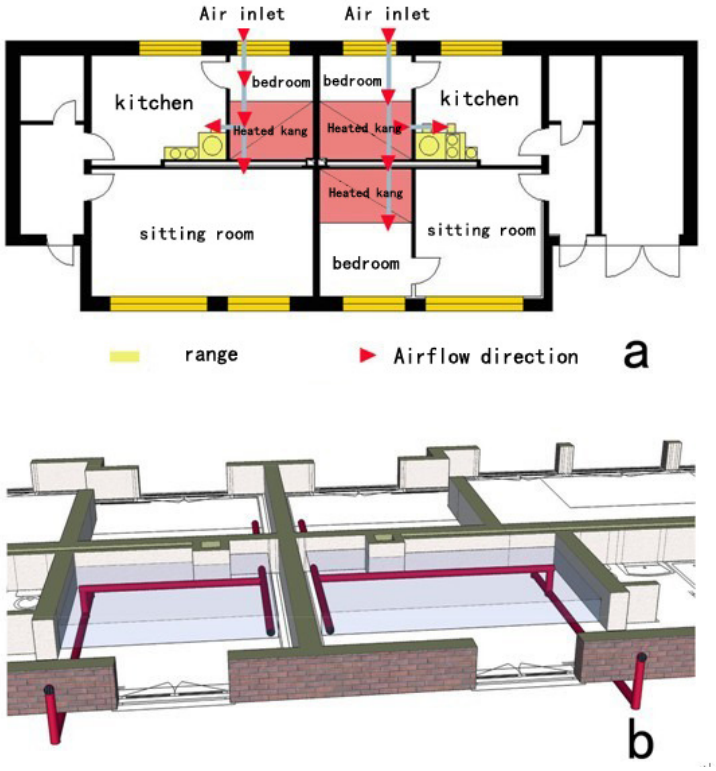


FIGURE 2: a: A Typical Residential Air Facilities Plan b: Fresh Air Facilities Isometric View.

2.3 Both Fresh Air Facilities into Problems

The author research the heilongjiang housing and urban and rural construction and French environment and energy control department design of heihe love bright outside the three DaoGou village close tree village residential use fire resistance heating and indoor air facilities in residential fire resistance, design considering the poor residents in use its plugging phenomenon (see chart 3). The reason is that outdoor cold air into the room temperature is too low, too fast, can't control, etc.

FIGURE 3: Typical Residential into Fresh Air Facilities Plugging Phenomenon.

3 FIRE RESISTANCE HEATING AND INDOOR AIR INTO INFRASTRUCTURE IMPROVEMENT STRATEGY

3.1 Fire Resistance Heating and Indoor Air into Infrastructure Improvement Goals

(1) Suitable for fresh air temperature

The goal is to improve in maintaining the characteristics of the system under the premise of, enables it to achieve the temperature requirements. To meet the People's Daily needs, improve the existing facilities for residents to bring the discomfort.

(2) The fresh air flow control

Indoor air supplements should have a certain amount of demand, but also can not be too big, To fresh flow control helps regulate the consumption of facilities, and the temperature of the guarantee.

3.2 Fire Resistance Heating and Indoor Air into Infrastructure Improvement Design

(1) Improve content

According to the problem of fresh air facilities reconstruction import mode and fire resistance, the combination of the way and the proposed control method.

(2) Improvement Principle

Based on the fire resistance heating, into the new facility design mode of action principle, in indoor and outdoor with a catheter into the fresh air for connection, and the conduit of the indoor air conveying all through the heated kang preheating, then through oven, kang and chimneys discharge outside. According to the wind tunnel to wind control principle set u-bend. The characteristics of rural house in northeast China, fresh air port work area cooling load is too large, only fresh air through the heat source can essentially solve the problem of fresh air to send. The study found that the disadvantages of fresh air in between each other is not isolated, such as supply air temperature difference, cold load, working blow a cold wind over, but restrict each other, related factors. Several parameters indirectly through the air supply outlet geometric characteristics (size, shape, tuyere type), resistance characteristics and installation parameters

(tuyere installation height, direction, etc.) reflected. The practical application can introduce rural house. In rural house can be through a facility, the outdoor air through the fire resistance heating systematically into indoor, which can overcome the cold region of the facility in a series of new problems. This facility model of the main technical measures is as follows:

(3) The fresh air tube U design

To effectively avoid the cold wind factors, into the fresh air into the FengKouChu device designed to U (see figure 4), concave to the outer wall. And the fresh air into the mouth at the entrance to the design window screen net block particles. Into the fresh air pipe buried under 900 mm.

(4) Coil heating

Increasing the length of the tube heated take, set to coil form. The ministry is divided into the fresh air part

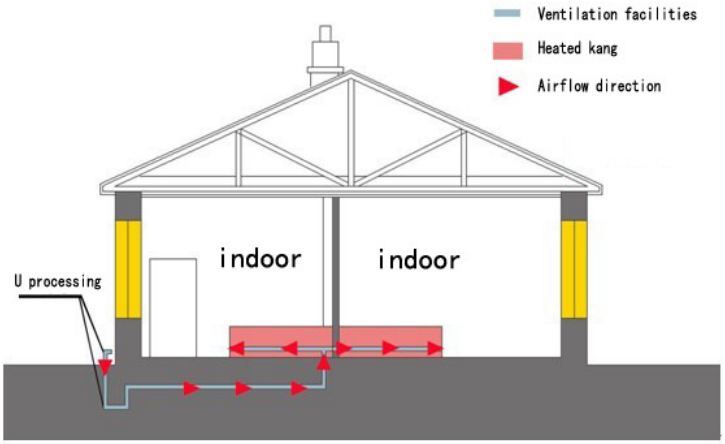


FIGURE 4: Typical Residential Ventilation Facilities u-shaped Ventilation Pipe Design.

of the power, temperature ascend, and integration of the important parts of the combination, so this part of the function is diverse, material also generally have all aspects of performance: reference standard JC646-1996 against odd strength index and inspection method has made the detailed provisions, the main consideration into fresh air pipe is thin wall (3 to 8 mm) fitting, and long-term heating use, service life requirements in 30 years, so the mechanical strength is a very important mechanical properties. Second fire resistance, heating pipe for a long time in heating environment use, its mechanical performance not significantly reduce called the nature of fire resistance. Heating pipe using the environment more complex, this requires heating pipe must have good fire resistance. Good toughness is into the fresh air in the pipe installation, maintenance of common problems. In this process, products must be from great impact and pressure, such as foot, hand to touch, the weight of the application, even by the strong impact and sharp (see chart 5).

(5) Facilities material

Into the fresh air at the entrance to new device with high thermal inertia of the materials, such as PVC, etc.; combined with

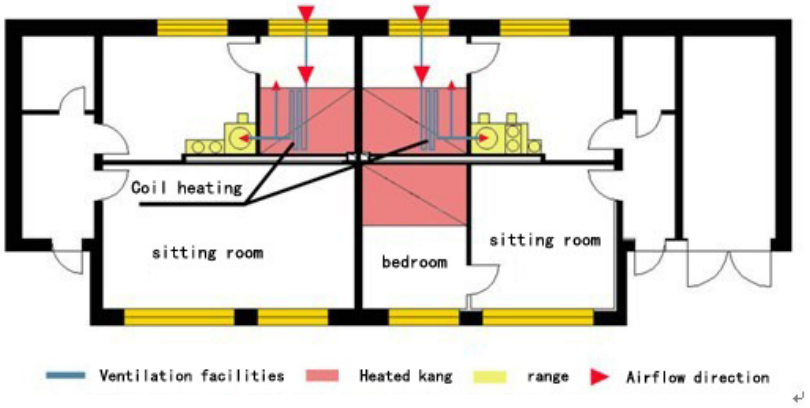


FIGURE 5: Typical Residential Ventilation Facilities Coil Design.

the fire resistance material selection of refractory material and prevent damage to the facilities.

(6) Proposed control valve

Control valve installed in the air port room side. But must function sensitive and with their into fresh air system mutual matching, should also be adaptable, easy to use, reliable and accurate, easy to maintain, durable. Among them, one of the key characteristics should be simple to use. Many controllers has many functions, but for town dwellers feel its operation is too complicated. Simple easy to use that can help people understand the operation of the control method and the fresh air system into an understanding.

4. FIRE RESISTANCE HEATING AND INDOOR AIR FACILITIES IN APPLIED RESEARCH

Due to the fire resistance of various shape, into the fresh air facilities in one of the application is also different, this paper illustrates two types of fire resistance application modes for practical reference.

(1) South kang application mode

Facilities system with this form of fire resistance with facilities (see chart 6 a) shows, air inlet design in side wall, the outlet were set up in the bedroom and kitchen, the corresponding setting control valve, to ensure that fresh air into the many and velocity, and the form of fire resistance combination should pay attention to air inlet aperture Settings, because in winter heating season leading northeast wind for the northwest wind, fresh air is with dust and pollutants, so attention should be paid to set up air inlet to ensure air quality.

(2) Full PuHang application modes

Facilities system with this form of fire resistance with facilities (see chart 6 b) shows, air inlet design in side wall, the outlet were set up in the bedroom and kitchen, the corresponding setting control valve, to ensure that fresh air into the many and velocity, because the shape of the heated kang area occupied the whole the size of the room, generally the air port Settings in the side wall of the inner edge, the arrangement of the time to try to consider side arrangement so that we can guarantee to avoid air mixing pollution, and the formation of uniform flow field, the bedroom air uniform flow to the sitting room, and then to the kitchen hearth crater, the chimney discharge.

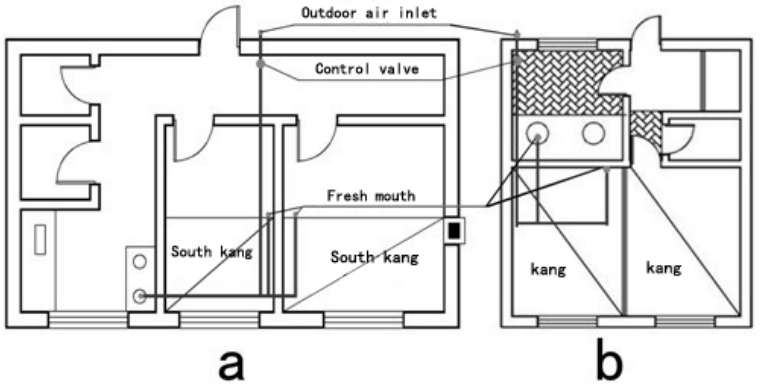


FIGURE 6: Into the Fresh Air Facilities in the Fire Resistance of The Application of The Housing.

5. CONCLUSION

(1) The investigation of northeast China's fire resistance heating room indoor new wind, detailed finishing the town heated kang residential air supplement mode, the typical case analysis of the current situation, discusses the new wind facilities both problems.

(2) Proposes fire resistance heating and indoor into fresh air facilities improvement strategies, including improvement goals and improved design scheme.

(3) Lists into new facilities and south kang, full PuHang these two types of fire resistance application mode.

6. REFERENCES

1. Hongguo Ren. Cold rural house fire resistance heating and ventilation integration research [j].2009.

2. Liu Man Liaoning province rural residential heating mode and energy research [j]. Building energy saving.2007.

3. Choulin Jiang. Tuyere is placed on displacement ventilation and tuyere selection experimental research Renewable energy.2007.

4. Pei Zhang. The cold region rural new solar heating technology design research [D].Shandong construction university.2007.

Arabidopsis thaliana Inspired Genetic Restoration Strategies

Donagh Hatton *donagh.hatton@nuim.ie*
Dept. of Computer Science
National University of Ireland, Maynooth
Maynooth, Ireland

Diarmuid P. O'Donoghue *diarmuid.odonoghue@nuim.ie*
Dept. of Computer Science
National University of Ireland, Maynooth
Maynooth, Ireland

Abstract

A controversial genetic restoration mechanism has been proposed for the model organism *Arabidopsis thaliana*. This theory proposes that genetic material from non-parental ancestors is used to restore genetic information that was inadvertently corrupted during reproduction. We evaluate the effectiveness of this strategy by adapting it to an evolutionary algorithm solving two distinct benchmark optimization problems. We compare the performance of the proposed strategy with a number of alternate strategies – including the Mendelian alternative. Included in this comparison are a number of biologically implausible templates that help elucidate likely reasons for the relative performance of the different templates. Results show that the proposed non-Mendelian restoration strategy is highly effective across the range of conditions investigated – significantly outperforming the Mendelian alternative in almost every situation.

Keywords: Evolutionary Algorithms, Genetic Restoration, Arabidopsis thaliana, Constrained Optimization

1. INTRODUCTION

The inspiration for this paper lies in a recently proposed genetic repair process in the model plant *Arabidopsis thaliana* [1]. *A. thaliana* is widely used in genetic studies, having a very fast life-cycle of around 6 weeks from germination to mature seed, making it ideal for longitudinal study and allowing comparison of multiple gnomes from the same species. It has a relatively short genome with about 157 million base-pairs encoding 27,000 genes and was the first plant to have its entire genome sequenced in the year 2000.

The reported study attests that some offspring of this plant inherit genetic material from individuals *other than* the direct parents. This controversial non-Mendelian hypothesis originated from studies of the HOTHEAD (*HTH*) gene, which impacts on formation of the epidermis and flower of this plant. The mutant *hth* form of the HOTHEAD gene is recessive and plants were studied that were homozygous for this recessive mutant allele (*hth/hth*). When these *hth/hth* mutants were allowed to self-fertilize, amazingly over 10% of the resulting progeny were of the normal wild type (*HTH/hth* or even *HTH/HTH*) - even though the normal *HTH* gene was not detected in either parent. This 10% rate of reversion is far higher than can be accounted for by random point mutations, which would generally occur with a frequency of the order of 1 per billion per allele per generation [2].

These findings are not consistent with the standard Mendelian model of inheritance and led to the controversial non-Mendelian theory. While these findings have been attributed to "pollen contamination" [3] [4], this claim itself has been contested [5].

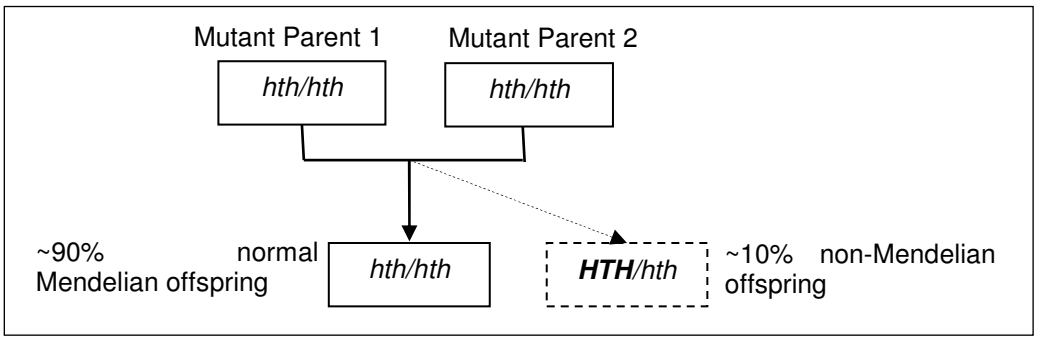


FIGURE 1: Approximately 10% of the offspring of hth/hth mutants revert back to the normal wildtype (either HTH/hth or even HTH/HTH). This finding is not consistent with normal Mendelian inheritance.

2. EVOLUTIONARY ALGORITHMS

Evolutionary algorithms, genetic algorithms [6] and related areas are based on Darwin's theory of natural selection [7]. They are widely used to find near optimal solutions to complex problems, wherein a population of (initially random) solutions are allowed to evolve under the guidance of a fitness function. Fitter genomes survive and reproduce more frequently, so that surviving individuals tend towards the optimal solution for a given problem.

While Evolutionary Algorithms (EA) are very effective and adaptable across a range of different problem types, they have a fundamental difficulty in dealing with constrained evolution. Constrained evolution addresses problems where the evolutionary process can generate genotypes that do not correspond to valid phenotypes. These infeasible solutions are inadvertently generated by either the process of reproduction or as a result of random mutations. These invalid (or infeasible) solutions are somewhat analogous to abortion events in natural evolution, where a viable phenotype cannot be produced from the genotype.

Some notable facets of our evolutionary strategy should be made clear at this point. First, our model is more obviously akin to the inheritance of sequences of genes, rather than the inheritance of lower-level base-pairs *per se*. Second, for our problems we can evaluate fitness directly from the genome sequence due to the simple and direct relationship between the genotype and phenotype for the investigated problems. Third, EA applied to permutation problems converge very slowly due to the irregular nature of the problem space. Finally, unlike the genetic restoration process proposed for *Arabidopsis thaliana*, we immediately repair genetic defects rather than repairing after a mutant solution creates a new solution.

2.1 Constrained Optimization Problems

Particularly challenging for EA are the permutation and combinatorial optimization problems, such as the Travelling Salesman's Problem (TSP) and Quadratic Assignment Problem (QAP). Solutions to these problems can be stated as identifying the optimal ordering of an initial list of items. The issue with employing an EA to solve such problems concerns ensuring that the core genetic operators of crossover and mutation do not introduce omission errors into the representation of solutions. Furthermore, due to the tightly constrained nature of these problems, the number of invalid solutions vastly out-numbers the valid solutions. Figure 2 demonstrates that as the size of the problem increases (across the horizontal axis), the invalid solutions vastly outnumber the valid solutions – note the exponential scale on the vertical axis. Thus, repairing these infeasible solutions is crucial to the effective performance of an evolutionary algorithm.

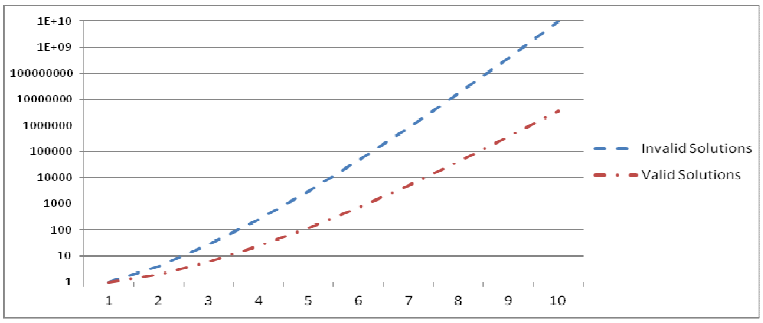


FIGURE 2: The Invalid Solutions Vastly Outnumber The Valid Solutions.

Before examining the restoration process, we briefly outline how genetic defects may be introduced into representations that have not been specifically designed for each problem type. For general representations of combinatorial optimization problems, invalid phenotypes can be easily identified by examining the genotype itself. Each duplicate gene denotes a genetic defect (see figure 3). In addition, these errors are accompanied by omission errors, as they occur simultaneously for this general representation scheme.

FIGURE 3: The processes of recombination/crossover (left) and mutation (right) easily generate invalid solutions to combinatorial optimization problems.

2.2 Template Driven Restoration in Evolutionary Optimization

Four distinct approaches have been adopted to enforce constraints on evolutionary searches [8] and this paper explores a novel strategy within the "genetic repair" approach. One approach to genetic repair has centred on the use of template solutions, which serve to restore genetic information that was accidentally removed during the reproduction or mutation phases of the algorithm. Template driven repair was introduced by Mitchell *et al.* [9], who compared the performance of static repair templates against a number of Mendelian-like templates. Later work showed that random repair templates [10] [11] generally produced better results, leading to their use in some standard application programs.

We now describe the template driven repair process as used for the combinatorial optimization problems. An erroneous genome is located by scanning its genes and locating any duplicate genes (we will return to the implicit dependency between the direction in which this process is conducted and the specific gene that is deemed to be the duplicate). Next, by comparison with the repair template we identify genes contained in the repair template, but which are not present in the invalid solution. These are highlighted in boldface on the lower line of figure 4. The restored solution in this case would be a b c d e f. However, if a different restoration template were used (say, a b e d c f) this would generate a different solution (a b e d c f). Thus, different repair templates have an impact on the fitness of the restored solutions. The rest of this paper concerns the impact that using different repair templates has on the quality of the solutions generated.

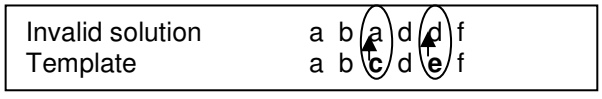


FIGURE 4: The repair template is used to restore information that was missing from the invalid solution.

The repair templates that are the focus of this paper are:

1. Parental templates. Repair templates are sourced from the parents of the invalid solutions. These restoration templates represent an alternative and more parsimonious Mendelian repair theory.

2. Ancestral templates. Templates are retrieved from ancient archived ancestors of the invalid solution. Repaired solutions are non-Mendelian as they contain genomic information from an individual other than the immediate parents. This paper investigates templates ranging from 10 generations to one thousand generations old.

3. Random templates. A new random template is generated for each invalid individual.

This paper explores the impact that a number of parameters have on the relative fitness of these genetic restoration templates. The main parameters investigated are problem type (TSP and QAP), problem size (ranging from 26 to 18152 genes in length), and the rate at which mutations impact on the population. We characterize this final parameter (mutation) into three distinct categories – low, tuned and high. Mutation rates have a very significant impact on solution quality, but manual tuning is generally required to find the optimal rate. In the remainder of this paper, the tuned mutation rate is near optimal. The low mutation rate conditions describe template performance below this rate while the high condition represents performance at mutation rates above the optimal. These three categories serve to adequately describe the (relative) performance of each of the different restoration templates. It should also be noted that optimal mutation rates vary widely between problem type, problem size, as well as other factors.

3. RESULTS

Results for two distinct problem types are presented in this section. The Traveling Salesman Problem (TSP) is the problem of visiting each of n cities once, with the goal of minimizing the total distance travelled. Solutions to the TSP are represented as a permutation of integers, representing the order in which cities are visited. The cost function for a TSP instance is the total route distance of the tour represented by a candidate solution. The TSP is a minimization problem, and so for all of the following results, a lower cost is better. The Quadratic Assignment Problem (QAP) is the problem of assigning *n* facilities to *n* different locations with the goal of minimizing the sum of the distances multiplied by a corresponding flow weights between the facilities. As with the TSP, a solution is represented as a permutation of integer values, representing the location of assignment of each facility in order. The QAP is also a minimization problem, so smaller values on the fitness axis are better. Problem instances were selected from the widely used TSPLIB [12] and QAPLIB [13] problem libraries. Specific benchmark problems are references, such as eil101. This is a 101 gene (city) instance of the TSP problem. Similarly, d18512 is a TSP problem whose solutions contain 18512 genes.

All of the experiments were conducted using tournament selection with a tournament size of 2, single point swap mutation and single point crossover. The experiments using eil101, nrw1379 and bur26a used a population of 100 while the 18,512 city problem used a population of 10. Mutation rate was selected to provide a rate that would be considered low, tuned and high according to problem size. All mutation rates discussed in the following section are expressed as a percentage of the genes in each solution, and represent the probability that a given gene will be mutated at each crossover event. Unless otherwise stated, mean values are derived from 10 separate experiments for each parameter combination. We compare ancestral templates from a range of ages. Due to the impact of the background mutation rate on these results, we compare performance for a variety of mutation rates, both above and below the optimum rate. The same genetic restoration process was used for both the TSP and QAP problems, indicating some generality for the approach adopted.

The key parameter that controls the age of the repair templates used is the Template Update rate (TU). TU controls the probability at each reproduction event that the current parent will be preserved as the ancestor to use for genetic restoration. For example, a TU of 1 means that the

parent of the new individual will be retained with probability p = 1, effectively resulting in the ancestor template always being constructed from the parent of the solution. A TU of 0.001 however, means that the ancestor template will be updated with that probability, extending the average age of the ancestor template to approximately 1000 generations.

3.1 Low Mutation Rates

The first set of results consists of the experiments conducted at rates of mutation lower than the optimal for problem type. In a typical EA, these conditions would result in insufficient mutation, reducing the speed of convergence of the population and impairing solution quality through a lack of diversity [6]. Figure 5 shows the results obtained for the 101 city TSP over 200,000 generations at the lower than optimal mutation rate of 0.01. A lower Template Update rate, corresponding to older ancestral templates, produced the best results (p-value < 0.01).

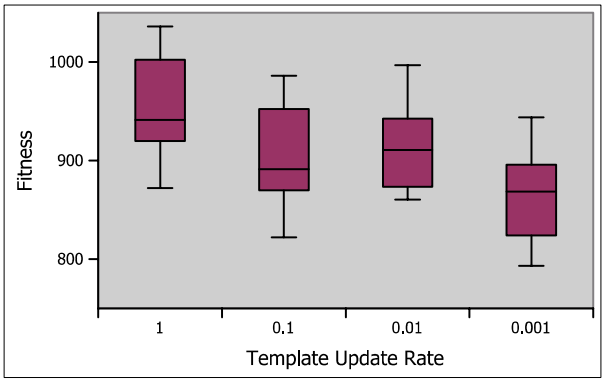


FIGURE 5: Effect of varying replacement rate on solution quality at a low mutation rate for TSP instance eil101.

Figure 6 shows the results obtained for the 1379 city TSP over 200,000 generations at the lower than optimal mutation rate of 0.01. Once again a lower Template Update rate produced the best results (p-value < 0.01), although the quality of results does at first decrease until the Template Update rate becomes lower and reverses that trend.

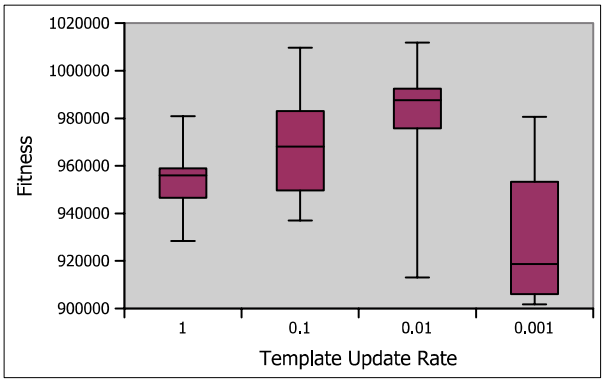


FIGURE 6: Effect of varying replacement rate on solution quality at a low mutation rate for TSP instance nrw1379.

Figure 7 shows the results obtained for the 18512 city TSP over 200,000 generations at the lower than optimal mutation rate of 0.001.

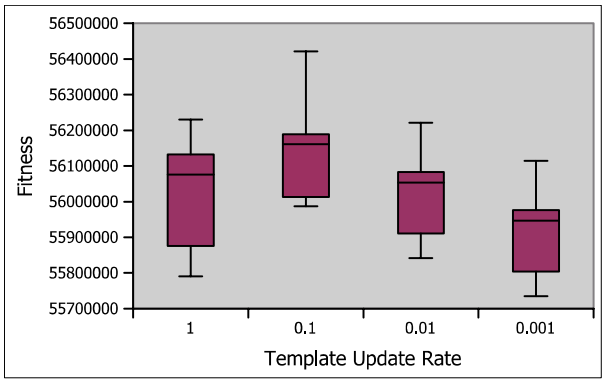


FIGURE 7: Effect of varying replacement rate on solution quality at a low mutation rate for TSP instance d18512.

Figure 8 shows the results obtained for the 26 facility QAP instance over 200,000 generations at the lower than optimal mutation rate of 0.01, although this time the difference in results is less strongly significant (p-value < 0.047)

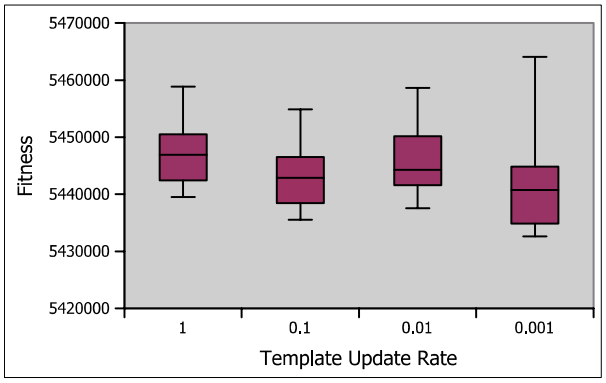


FIGURE 8: Effect of varying replacement rate on solution quality at a low mutation rate for QAP instance bur26a.

In all of the 4 problem instances investigated, a lower Template Update rate produced solutions with a better overall fitness cost. This indicates that using older ancestral templates to restore genetic defects in conditions of lower than optimal mutation frequency results in better convergence towards the optimum. In the case of the bur26a QAP problem, the upper and lower bound of the best result set were equivalent to the upper and lower bound of all the results obtained for that problem set. This indicates that the use of older ancestral templates can serve to introduce much greater genetic diversity into the population in the absence of sufficient diversity via mutation.

3.2 High Mutation Rates

This set of results consists of the experiments conducted at rates of mutation higher than the optimal for problem type. In a typical EA, these conditions would result in too much mutation, reducing the speed of convergence of the population and impairing solution quality [6]. Figure 9 shows the results obtained for the 101 city TSP over 200,000 generations at the higher than optimal mutation rate of 0.5. As with the experiments at a lower Mutation Rate, a lower Template Update rate produced the best results (p-value < 0.01).

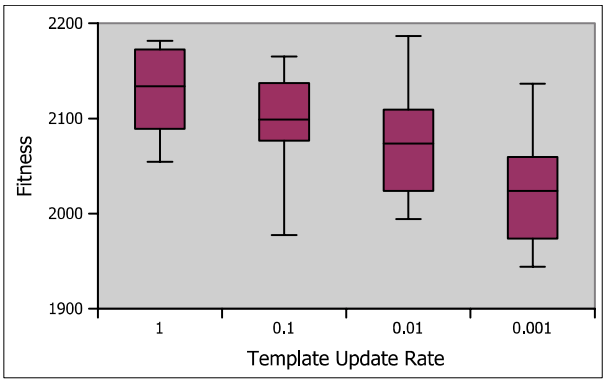


FIGURE 9: Effect of varying replacement rate on solution quality at a high mutation rate for TSP instance eil101.

Figure 10 shows the results obtained for the 1379 city TSP over 200,000 generations at the higher than optimal mutation rate of 0.5. The experiments using the oldest ancestral templates produced the best results (p-value < 0.01).

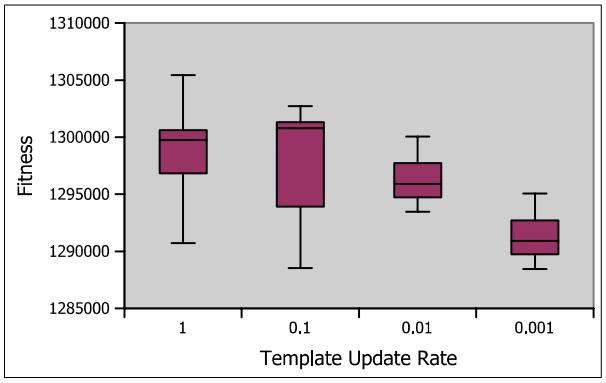


FIGURE 10: Effect of varying replacement rate on solution quality at a high mutation rate for TSP instance nrw1379.

Figure 11 shows the results obtained for the 18512 city TSP over 200,000 generations at the higher than optimal mutation rate of 0.5. The experiments using the oldest ancestral templates produced the best results (p-value < 0.01).

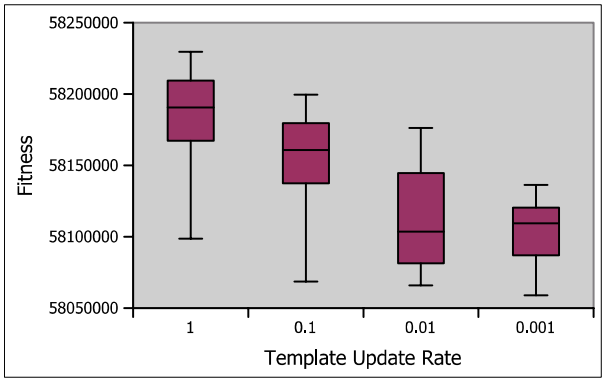


FIGURE 11: Effect of varying replacement rate on solution quality at a high mutation rate for TSP instance d18512.

Figure 12 shows the results obtained for the 26 facility QAP instance over 200,000 generations at the higher than optimal mutation rate of 0.5. In this case the difference between parameters was not enough to be statistically significant. This may be due to the excessively negative effect that a high mutation rate can have on a problem as heavily constrained as the QAP.

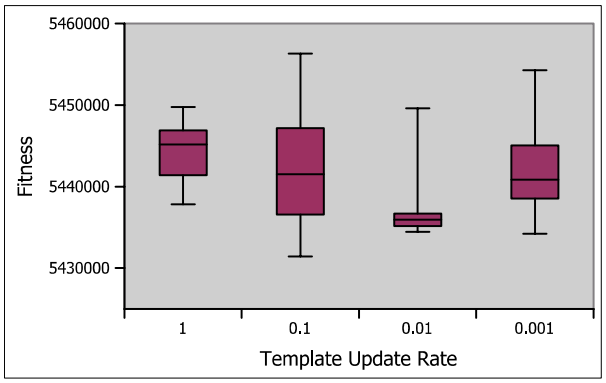


FIGURE 12: Effect of varying replacement rate on solution quality at a high mutation rate for QAP instance bur26a.

In 3 of the 4 problem instances investigated, a lower Template Update rate produced solutions with a better overall fitness cost. This indicates that using older ancestral templates to restore genetic defects in conditions of higher than optimal mutation frequency also results in better convergence towards the optimum.

3.3 Tuned Mutation Rates

This set of results consists of the experiments conducted at rates of mutation that were manually tuned per problem to produce stronger results. Typically, this approach can take a long time and is error prone [14]. For the TSPLIB and QAPLIB libraries, this is an easier task, as these particular problem instances have been extensively studied. Ordinarily, a tuned mutation rate would result in an optimal amount of mutation, maximizing the solution quality produced and minimizing the time required to arrive at a solution. Figure 13 shows the results obtained for the 101 city TSP over 200,000 generations at the tuned mutation rate of 0.1. In this situation there is no statistically significant difference between the test cases.

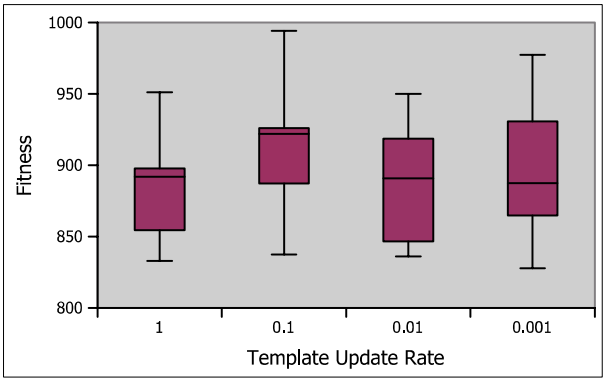


FIGURE 13: Effect of varying replacement rate on solution quality at a tuned mutation rate for TSP instance eil101.

Figure 14 shows the results obtained for the 1379 city TSP over 200,000 generations at the tuned mutation rate of 0.1.

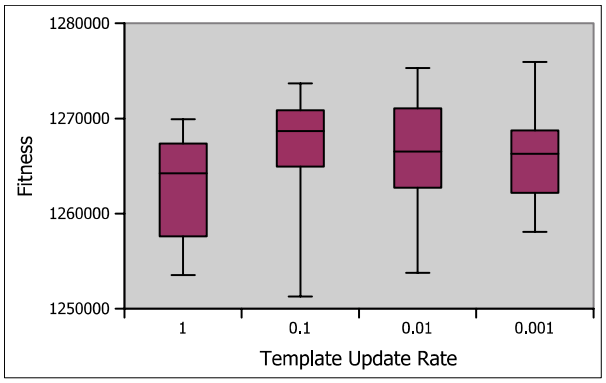


FIGURE 14: Effect of varying replacement rate on solution quality at a tuned mutation rate for TSP instance nrw1379.

Figure 15 shows the results obtained for the 1379 city TSP over 200,000 generations at the tuned mutation rate of 0.1.

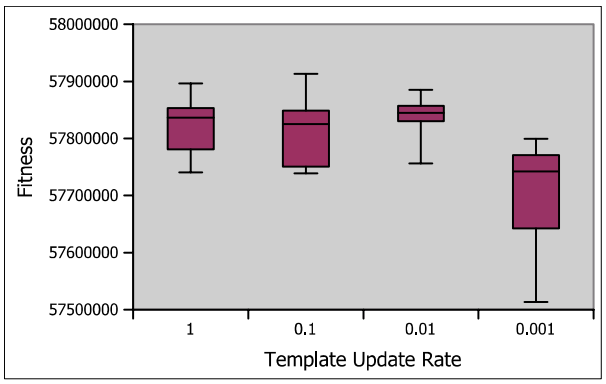


FIGURE 15: Effect of varying replacement rate on solution quality at a tuned mutation rate of 0.01 for TSP instance d18512.

Figure 16 shows the results obtained for the 26 facility QAP instance over 200,000 generations at the tuned mutation rate of 0.1. In this case the lower template update rate corresponding to older ancestral templates produced the best results.

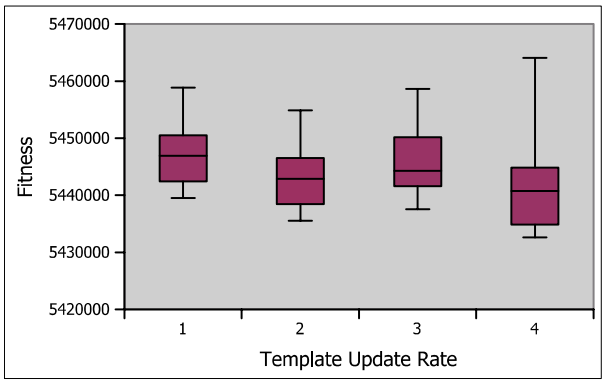


FIGURE 16: Effect of varying replacement rate on solution quality at a mutation rate of 0.1 for QAP instance bur26a.

3.4 Diachronic Analysis of Convergence

Figure 17 shows the results of a diachronic analysis of the improvement of solution quality for the eil101 TSP instance. Corresponding with figure 5, the lowest Template Update rate and therefore the oldest non-Mendelian ancestral templates are producing solutions that reach a better local optimum in fewer generations.

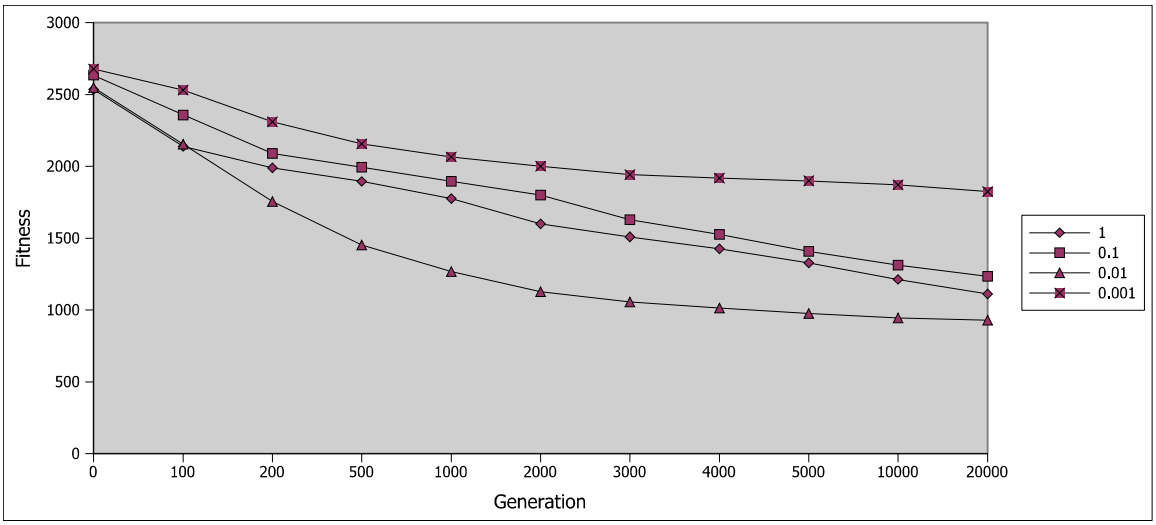


FIGURE 17: Rates of convergence of solutions for 101 city TSP instance by Template Update Rate.

3.5 Retention Policy and Repair Direction

In addition to a series of experiments in which the Mutation Rate and Template Update rate were varied, we evaluated the effect of using a different template retention policy and repair direction on the quality of results obtained. The template retention policy refers to the criteria that are used to select which template out of a possible set should be used for the purposes of Genetic Restoration. The policies investigated were; Best, Worst and Random. With Best, the ancestor with the lowest and therefore best fitness cost was selected to be used as the template in the event that an invalid solution was generated. With Worst, the ancestor with the highest and therefore worst fitness cost was selected to be used as the template and with Random the choice of policy was split between Best and Worst with equal probability of either occurring. Figure 18 gives the results of this comparison, using solutions generated for the nrw1379 TSP instance. There is no statistical significance attributed to deciding to use one policy over another.

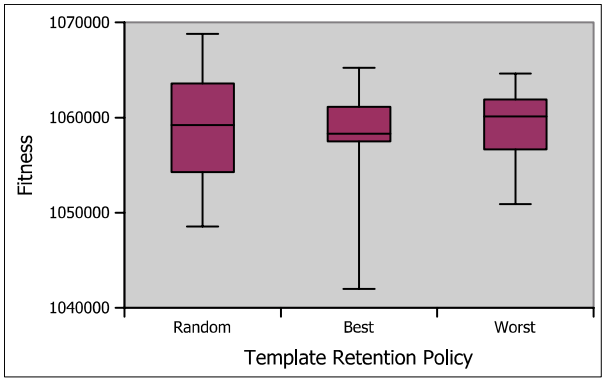


FIGURE 18: Effect of varying template retention policy on solution quality.

Repair direction refers to the order in which repair operations are carried out during the restoration phase. There are 3 choices for repair direction; Left to Right, Right to Left and

Random. The difference between these choices refers to the order in which the repair template is read. Left to Right will scan the repair template from left to right, and if it finds a valid replacement gene it will replace the defective gene with it. Right to Left scans the template in the opposite order. Defective genes are always replaced from left to right; the repair direction refers only to the order in which the repair template is scanned. With a Random strategy the choice of direction is split equally between Left to Right and Right to Left. It is important to note that the difference in repair direction only has an impact in situations where more than 2 genes in the solution are defective, as if only 1 gene is problematic, the order in which repair occurs will have no impact on the choice of replacement gene. Figure 19 shows the results of varying repair direction for instances of the nrw1379 problem, the Random strategy producing the best results with $p < 0.01$. It is likely that the Random strategy is more effective due to the additional diversity introduced to the population by varying repair direction.

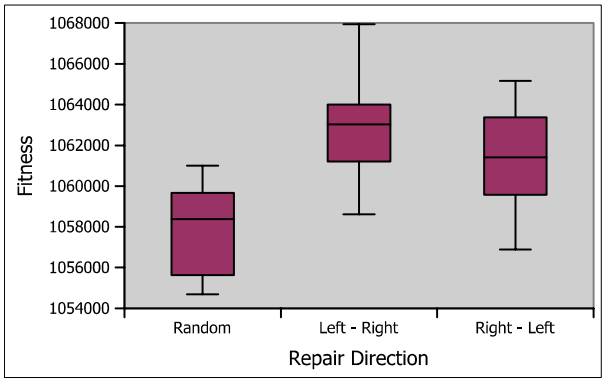


FIGURE 19: Effect of varying repair direction on solution quality.

3.6 Comparison to a Random Template

One alternate approach that has been used in the literature for sourcing repair templates is that of generating a randomized genetic sequence to use as repair data [10]. Figure 20 gives results comparing the effectiveness of random template generation as opposed to ancestral repair templates at low, tuned and high mutation rates respectively. For both low and high mutation rates, the ancestral template update policy produced fitter solutions (p-value < 0.001). At the tuned mutation rate the best results were still obtained with the ancestral repair template but the results were less statistically significant (p-value < 0.05).

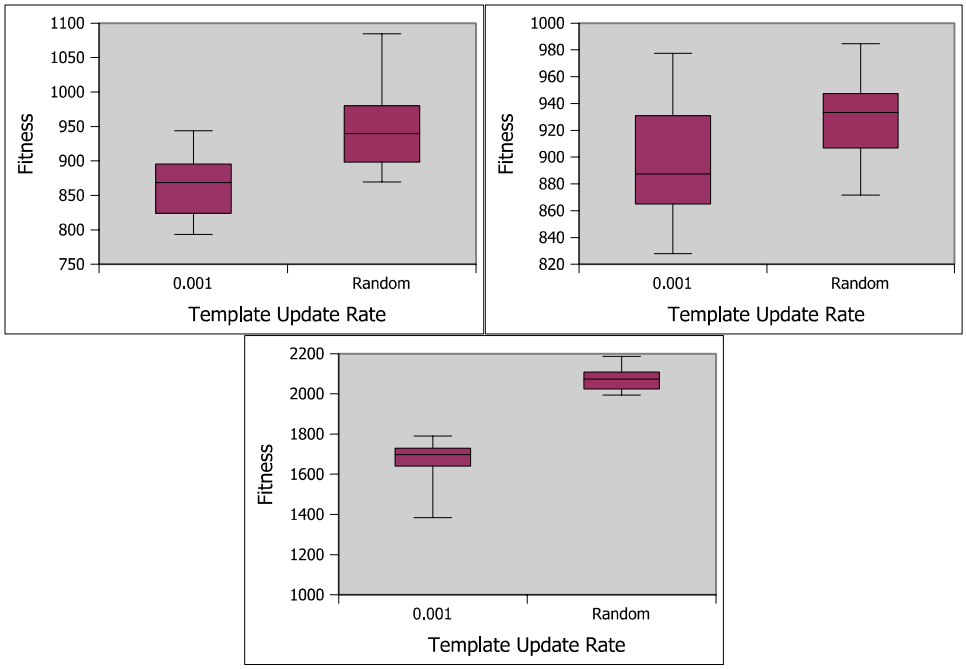


FIGURE 20: Comparison of ancestral template repair vs. random template repair across instances of the 101 city TSP problem at low (top left), tuned (top right) and high (bottom) mutation rates.

3.7 Analysis of Results

The results obtained at rates of mutation outside a tuned rate display a clear pattern. With both lower and higher mutation rates, older non-Mendelian repair templates produced the best results in each experiment. Outside of a small zone where algorithmic parameters such as mutation rate are manually tuned for optimal results, the Mendelian repair template either fails to perform, or performs no better than non-Mendelian ancestral repair. Even within optimal parameter configurations the non-Mendelian approach produces results as good as the Mendelian. We hypothesize that older templates produce fitter solutions than more recent templates for parameter combinations outside of this tuned zone because they exert an implicit clamping factor on the amount of mutations that can occur. In the presence of low mutation rates, the diversity in a population is limited. In practice, this means that the likelihood of ancestral templates introducing diversity is heavily influenced by their genomic distance from the candidate solution. In the presence of high mutation rates, the size of the search space examined is increased, as mutations that result in larger genomic changes are more likely. In these instances, repair is required more often, as shown by Mitchell et al. [9]. In this situation, older ancestral templates are able to correct the genetic defects which appear in greater amounts, while still retaining enough diversity to allow effective exploration of the search space.

3.8 Discussion

Genetic restoration appears to offer a new approach to managing population diversity by leveraging the contents of archived individuals. Ancestral genes are introduced into the current population to increase diversity, with the introduced genes being combined with genes from the current population. Introducing ancestral genes ensures that the included information is reasonably fit, avoiding the penalty associated with introducing true randomness. For complex optimization problem such as those discussed in this paper, we can think of genetic restoration as being somewhat akin to backtracking in traditional search.

Local search algorithms typically use backtracking to overcome local minima and reach the global maximum. An EA implicitly performs this sort of backtracking through the process of generating additional candidate solutions that converge in parallel [6]. Genetic restoration introduces an additional level of backtracking at the level of individual genes rather than at the level of the entire

genome. Implicit backtracking is facilitated by the re-introduction of genes originating from ancestors of the current population – creating mosaic individuals that combine current genes with some ancient genes. As far as the authors are aware this is the first use of ancient ancestral template repair in evolutionary computation, building on [15].

Further research possibilities for ancestral genetic repair include investigating its use to optimize real valued functions as well as deriving alternate approaches to other population based optimization methods such as differential evolution or simulated annealing. Additionally, ancient repair templates show good performance across a wide range of mutation rates, indicating that they may have a potential application in automatic parameter tuning.

4. CONCLUSIONS

In 2005, Lolle et al. published work [1] that indicated the possible existence of non-Mendelian inheritance in Arabidopsis thaliana. This repair mechanism makes use of ancestral genetic information to repair genetic errors in individual plants. We investigate an application of this strategy to handling constraint violations in an Evolutionary Algorithm (EA). In our EA, each individual records a collection of ancestral templates for the purposes of correcting genetic defects in the form of invalid candidate solutions. Results are presented relating to three instances of the Traveling Salesman's Problem (TSP) and one instance of the Quadratic Assignment Problem (QAP).

Some general observations can be made across the results. First, at relatively low rates of background mutation, where a typical EA suffers with stagnation due to a lack of genetic diversity, the best results were consistently produced by the most ancient non-Mendelian repair templates (around 1000 generations old). Second, at relatively high rates of background mutation, where a typical EA suffers from too much diversity, the best results were again produced by older repair templates. Third, when mutation rates were tuned by hand to provide a good rate of convergence for the problem instances examined, older non-Mendelian repair templates produced results that were as good as the Mendelian repair templates.

Overall our results indicate that older (and thus non-Mendelian) ancestral repair templates provide a better general-purpose repair mechanism than the Mendelian alternative. On problems for which optimal mutation rates are unknown, using ancient repair templates produces reliably better results. It appears that the strategy proposed in [1] is surprisingly effective across a range of problems and conditions.

5. REFERENCES

1. S. J. Lolle, J. Victor, J. Young, and R. Pruitt, "Genome-wide non-mendelian inheritance of extra-genomic information in arabidopsis," *Nature*, vol. 434(1), pp. 505–509, 2005.

2. D. Weigel and G. Jurgens, "Genetics: Hotheaded healer," *Nature*, vol. 434(443), pp. 443–443, 2005.

3. P. Peng, S. Chan, G. Shah and S. Jacobsen, "Increased outcrossing in hothead mutants." *Nature*, vol. 443, pp. E8–E8, 2006.

4. R. Mercier, S. Jolivet, J. Vignard, S. Durand, J. Drouaud, G. Pelletier, and F. Nogue, "Outcrossing as an explanation of the apparent unconventional genetic behavior of arabidopsis thaliana hth mutants," *Genetics*, vol. 180(4), pp. 2295–2297, Dec. 2008.

5. M.T. Hopkins, A.M. Khalid, P.C. Chang *et al.* (2013) "*De novo* genetic variation revealed in somatic sectors of single Arabidopsis plants", *F1000Research* vol. 2 no. 5, 2013,

6. D. E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", 1st ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1989.

7. C. Darwin, "On the origin of species", New York :D. Appleton and Co, 1861

8. S. Salcedo-Sanz, "A survey of repair methods used as constraint handling techniques in evolutionary algorithms," *Computer Science Review*, vol. 3, pp. 175–192, 2009.

9. G. G. Mitchell, D. O'Donoghue, and A. Trenaman, "A new operator for efficient evolutionary solutions to the travelling salesman problem," in *Applied Informatics*, 2000.

10. D. Lichtblau, "Discrete optimization using Mathematica," in *Proceedings of the World Conference on Systemics, Cybernetics, and Informatics (SCI 2002)*, vol. 16, 2002.

11. G. G. Mitchell, D. P. O'Donoghue, D. Barnes, and M. McCarville, "Generepair - a repair operator for genetic algorithms," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO)*, 2003, pp. 235–239.

12. G. Reinelt, "Tsplib - a travelling salesman problem library," *ORSA Journal of Computing*, vol. 3, pp. 376–384, 1991.

13. R. E. Burkard, S. E. Karisch, and F. Rendl, "Qaplib - a quadratic assignment problem library," *J. of Global Optimization*, vol. 10(4), pp. 391–403, Jun. 1997.

14. Z. Michalewicz and M. Schmidt, "Evolutionary algorithms and constrained optimization," in *Evolutionary Optimization*, R. Sarker, M. Mohammadian, and X. Yao, Eds. New York: Kluwer Academic Publishers, pp. 57–86, Feb. 2002, .

15. A. Fitzgerald and D. P. O'Donoghue, "Genetic repair for optimization under constraints inspired by Arabidopsis thaliana," in *Proceedings of the 10th International Conference on Parallel Problem Solving from Nature (PPSN)*, pp. 399–408, 2008,.

Using Brain Waves as New Biometric Feature for Authenticating a Computer User in Real-Time

Kusuma Mohanchandra *kusumalak@gmail.com*
Associate Professor/Department of Computer Science & Engineering
Dayananda Sagar of Engineering
Bangalore, 560078, India

Lingaraju G M *gmlraju@gmail.com*
Professor/Department of Information Science & Engineering
M S Ramaiah Institute of Technology
Bangalore, 560054, India

Prashanth Kambli *prash.kamblil@gmail.com*
Assistant Professor/Department of Information Science & Engineering
M S Ramaiah Institute of Technology
Bangalore, 560054, India

Vinay Krishnamurthy *vinayk.url@gmail.com*
Student, Department of Computer Science
Stony Brook University
Stony Brook - 11790, NY, USA

Abstract

In this paper we propose an Electroencephalogram based Brain Computer Interface as a new modality for Person Authentication and develop a screen lock application that will lock and unlock the computer screen at the users will. The brain waves of the person, recorded in real time are used as password to unlock the screen. Data fusion from 14 sensors of the Emotiv headset is done to enhance the signal features. The power spectral density of the intermingle signals is computed. The channel spectral power in the frequency band of alpha, beta and gamma is used in the classification task. A two stage checking is done to authenticate the user. A proximity value of 0.78 and above is considered a good match. The percentage of accuracy in classification is found to be good. The essence of this work is that the authentication is done in real time based on the meditation task and no external stimulus is used.

Keywords: Cognitive Biometrics, Authentication, Brain Computer Interface, Electroencephalogram, Power Spectral Density.

1. INTRODUCTION

In this computer driven era, with the increase in security threats, securing and managing the resources has become a more complex challenge. Maintaining and managing access while protecting the user's identity and computer resources has become increasingly difficult. Therefore, it is crucial to design a high security system that has a strong authentication process to authenticate an individual. Authentication, verifying the user, who he claims to be, is the central to all security systems. With the world getting ready to transit from Graphic User Interface (GUI) to Natural User Interface (NUI) technology, where it is possible to communicate with computers by using touch, gestures, voice, expressions, emotions and thoughts. In this context, we have made an attempt to build an authentication system based on thoughts.

The common authentication approaches are those based on personal identification number (PIN) and password. However, these can be easily compromised by methods such as 'shoulder surfing' [1]. The biometric approaches based on the biological characteristics of humans have distinct advantages over traditional methods, as they cannot be hacked, stolen or transferred from one person to another as they are unique for each person. But, as the biological characteristic of a person change with time and age, it is required to find an alternative biometric trait that can distinguish between individuals. Multimodal fusion for identity verification [2] has shown great improvement compared to unimodal algorithms where they propose to integrate confidence measures during the fusion process. These methods are used either to enhance the performance of a multimodal fusion algorithm or to obtain a confidence level on the decisions taken by the system.

Existing technologies mostly use fingerprints, speech, facial features, iris and signatures as a base for an authentication or an identification system. These traits however, are known to be vulnerable to falsification as it is possible to forge or steal. Therefore, new types of physiological features that are unique and cannot be replicated are proposed [3] for an identification system. This paper focuses its attention to the electroencephalogram (EEG) signal as a biometric. The EEG based biometrics is widely being considered in security sensitive areas like banks, labs and identification of criminal in forensic. It can be used as a component of National e-identity card in government sector, as they have proven to be unique between people.

Brain-computer interface (BCI) is an emerging technology which aims to convey people's intentions to the outside world directly from their thoughts, enhancing cognitive capabilities and is a direct communication pathway between a brain and an external device. A common method for designing BCI is to use EEG signals extracted during mental tasks [4]. EEG is the neurophysiological measurement of electrical activity in the brain recorded by scalp electrodes (sensors) and represents a summation of post-synaptic potentials from a large number of neurons. Studies have shown that Brain wave pattern for each individual is unique and thus can be used for biometric purpose. EEG-based biometry [5] is an emerging research topic. Very little work has been done in this area, focusing more on person identification than person authentication. Person authentication aims to accept or reject a person claiming an identity, i.e. comparing a biometric data to one template, while the goal of person identification is to match the biometric data against all the records in a database [6]. In our work, we have made an attempt to authenticate a system, rather than identification. EEG is used to extract reliable features of brain signals [7]. Brain waves measured by EEG represent a summary of brain electrical activity at a recording point on the scalp i.e. the fusion of delta, theta, alpha/mu, beta and gamma waves in frequency band.

FIGURE 1: User Wearing the Emotiv Epoc Headset.

In this work, we investigate the use of brain activity for person authentication. It has been studied that the brain-wave pattern of every individual is unique [8] and the signals captured through the EEG can be used for biometric authentication. Person authentication aims to accept or reject a

person claiming an identity. An EEG EPOC headset with 14 channels manufactured by Emotiv Inc. is used for signal acquisition. The data acquired from these multi sensors are coordinated and managed to give the desired performance. N. Xiong et al [9], presents a comprehensive review of multi-sensor management in relation to multi-sensor information fusion, describing its place and role in the larger context, generalizing main problems from existing application needs, and highlighting problem solving methodologies. The purpose of data fusion [10] is to produce an improved model or estimate of a system from a set of independent data sources.

We perform data acquisition, feature extraction, matching the feature vector with the stored template all in real time. As data from multiple channels is fusion, we have used Power Spectral Density as a reliable feature [11]. Hence Power Spectral Density is used as the key feature in this work. After obtaining the features, Principal Component Analysis (PCA) is performed to obtain relevant features from the high dimensional data. The obtained feature vector is then compared against a previously stored feature vector for the same person using template matching. A two stage checking is done to authenticate the user. A single biometric with multiple matches [12] is considered. The match is considered good if the result of the comparison is greater than the threshold value which has been set to 0.78 after repeated trials keeping in mind the need to satisfy low False Acceptance Error (FAE) and False Rejection Error (FRE). The decision threshold of a system is set so that the proportion of false rejections will be approximately equal to the proportion of false acceptances called as Equal Error Rate [6]. We have developed a GUI, to let a user lock his computer screen when required and unlock the same by recording his brain activity (EEG signals) as a password for the system. This authentication system was successfully demonstrated as a pilot project and proof of concept [13].

An identity authentication system has to deal with two kinds of events: either the person claiming a given identity is the one who he claims to be (in which case, he is called a *client*), or he is not (in which case, he is called an *impostor*). Moreover, the system may generally take two decisions: either *accept* the *client* or *reject* him and decide he is an *impostor* [6]. The main aim is to keep the False Acceptance Error (FAE) and the False Rejection Error (FRE) close to zero. The Education Edition SDK by Emotiv Systems includes a research headset: a (plus CMS/DRL references, P3/P4 locations) high resolution, neuro-signal acquisition by wireless sensors and processing wireless neuroheadset. Channel names based on the International 10-20 locations are: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 and AF4. The Education Edition SDK [14] also includes a proprietary software toolkit that exposes the APIs and detection libraries. The SDK provides an effective development environment that integrates well with new and existing frameworks.

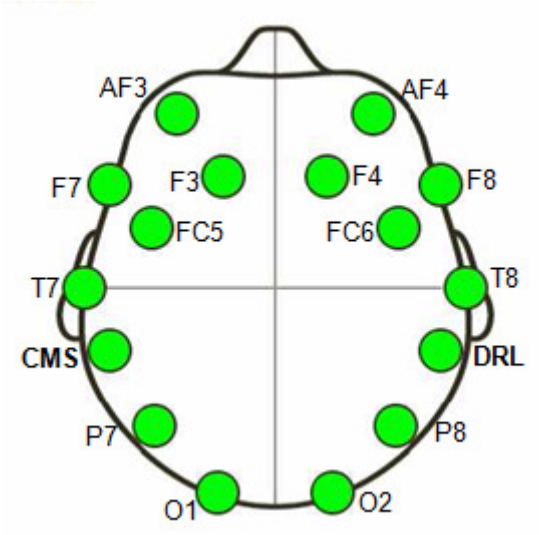


FIGURE 2: Illustration of Location of Electrodes on the Emotiv Headset [14].

2. RELATED WORK

EEG based person authentication was first proposed by Marcel [6]. They proposed the use of Power Spectral Density as the feature, and a statistical framework based on Gaussian Mixture Models (GMM) and Maximum A Posteriori Model (MAP) Adaptation on speaker and face authentication. The potential of their method is shown by simulations using strict train/test protocols and results. Person identification based on spectral information [15] extracted from the EEG is addressed by M. Poulos, et al, . Neural network classification was performed on real EEG data of healthy individuals to experimentally investigate the connection between a person's EEG and genetically specific information. The proposed method has yielded correct classification scores in the range of 80% to 100%, showing evidence that the EEG carries genetic information for person identification.

A novel two-stage biometric authentication [1] method was proposed by Ramaswamy Palaniappan. The feature extraction methodology includes both linear and nonlinear measures to give improved accuracy. Their results show that the combination of two-stage authentication with EEG features has good potential as a biometric as it is highly resistant to fraud. Principal Component Analysis (PCA) is used for dimension reduction of the feature vector keeping only the most discriminatory features, as the features have a high degree of redundancy.

3. METHODOLOGY

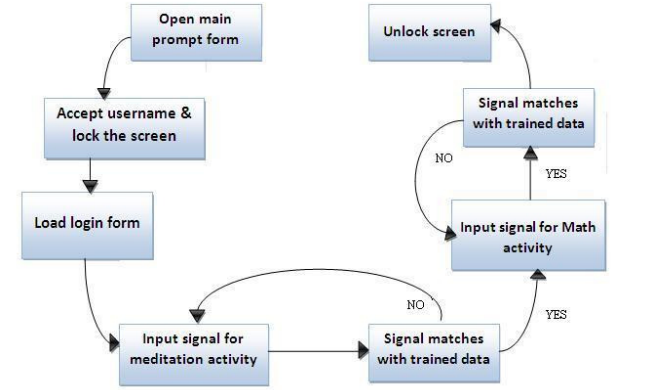


FIGURE 3: Framework of the Model.

A conceptual framework of the present work is shown in figure 3.

3.1 Data Acquisition

EEG signals are recorded with the Emotiv EPOC headset which uses 14 integrated sensors located at standard positions of the International 10-20 system (Fig: 2). Sensors are placed on the scalp using a conductive gel, after preparing the scalp area by light abrasion to reduce electrode-scalp impedance. The sampling rate is 128Hz [16]. The total time of each recording is 10 seconds. The subject is instructed to avoid blinking or moving his body during the data collection to prevent the noise caused due to artifacts [17]. So, no artifact rejection or correction is employed. Artifacts due to eye blinks produces a high amplitude signal called Electrooculogram (EOG) that can be many times greater than the EEG signals required by us [18]. The dataset from normal subjects are recorded for two active cognitive tasks during each recording session.

- *Meditation activity:* The subject is asked to meditate for a fixed period of time while his brain waves are recorded.

- *Math activity:* The subject is given non-trivial multiplication problems, such as 79 times 56 and is asked to solve them without vocalizing or making any other physical movements. The problems were designed so that they could not be solved in the time allowed [19].

3.2 Preprocessing and Feature Extraction

The EEG data is segmented. Channel spectral power for three spectral bands Alpha, Beta and Gamma is computed. 14 x 3 = 42 features are extracted for each segment of the data. PCA is applied to reduce the feature size. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible [20]. We have considered only those components that contribute to 85% (this value has been chosen after repeated trials) of the total variance for signal matching. The power spectral density (PSD) reflects the 'frequency content' of the signal or the distribution of signal power over frequency [21]. PSD is a positive real function of a frequency variable associated with a stationary stochastic process. It is the measure of the power strength at each frequency. In other words, it shows at which frequencies variations are strong and at which frequencies variations are weak [18]. The unit of PSD is energy per frequency (width). Computation of PSD can be done directly by the method of Fourier analysis or computing auto-correlation function and then transforming it.
The Discrete Fourier transform is given by

$$X(f) = \sum_{i=1}^{N} x(i) w_N (i-1),$$

Where

$$w_N = \exp(2\pi i) / N,$$

is the Nth root of unity. Power spectral density is given by

$$S_x(f) = \frac{1}{N} \sum_{i=1}^{N} | X(f) |^2$$

The channel spectral power is the measure of the total power between two frequencies and is given by:

$$P_{f1,f2} = \int_{f1}^{f2} S_x(f) df,$$

where (f1, f2) is the frequency band and $S_x(f)$ is the power spectral density. The inter-hemispheric channel spectral power differences in each spectral band are given by P_{diff} = (P1 – P2) / (P1 + P2) where P1 and P2 are the powers in different channels in the same spectral band but in the opposite hemispheres.

3.3 Classification

The obtained feature vector is compared against a previously stored feature vector for that subject, using Euclidean Distance for template matching. The match is considered good if the result of the comparison is greater than the threshold value which has been set to 0.78 after repeated trials keeping in mind the need to satisfy low False Acceptance Error (FAE) and False Rejection Error (FRE). A proximity value of 0.78 and above is considered a good match.

3.4 Implementation

The authentication system was realized by developing an application which would lock and unlock the screen. Initially the screen is locked and a subject's EEG signals for two mental tasks

are recorded and stored as a reference, called the training phase. If the screen is to be unlocked, the subject's brain waves are recorded again and matched with the earlier stored sample. If there is a considerable match, then the screen is unlocked, otherwise it will stay locked. The description of the working prototype is outlined as:

- *Training of the system:* The brain waves of the user are recorded when he performs the mental tasks such as meditation and math activity.
- *Feature extraction:* The channel spectral power in the alpha, beta and gamma spectral bands of both the mental tasks is computed. Feature reduction technique is applied, to reduce the dimension of the features.
- *Creating user profile:* These features are stored in a separate file as the user's profile.
- *Authenticating:* The brain waves of the person are recorded in real time for the same set of activities as in the training. Features are extracted from these recorded waves. Feature reduction is performed using Principal Component Analysis and these features are matched with the previously stored features. The feature extraction and matching part are coded in MATLAB, while the UI part is designed and coded in C#. The MATLAB codes are converted to Common Language Runtime (CLR) compliant library (*.dll file). These files are then referenced in C# by means of the *using* statement and adding an appropriate reference.

The User Interface diagram (Fig 4) explains the various stages and steps involved from the user's perspective. It depicts the different forms involved in the application for user interface. The following steps act as a walkthrough for the application.
Step 1: The initial screen which is the main prompt screen (Fig 5) facilitates the user to perform the lock screen, add/remove user, change account name and restore activities.

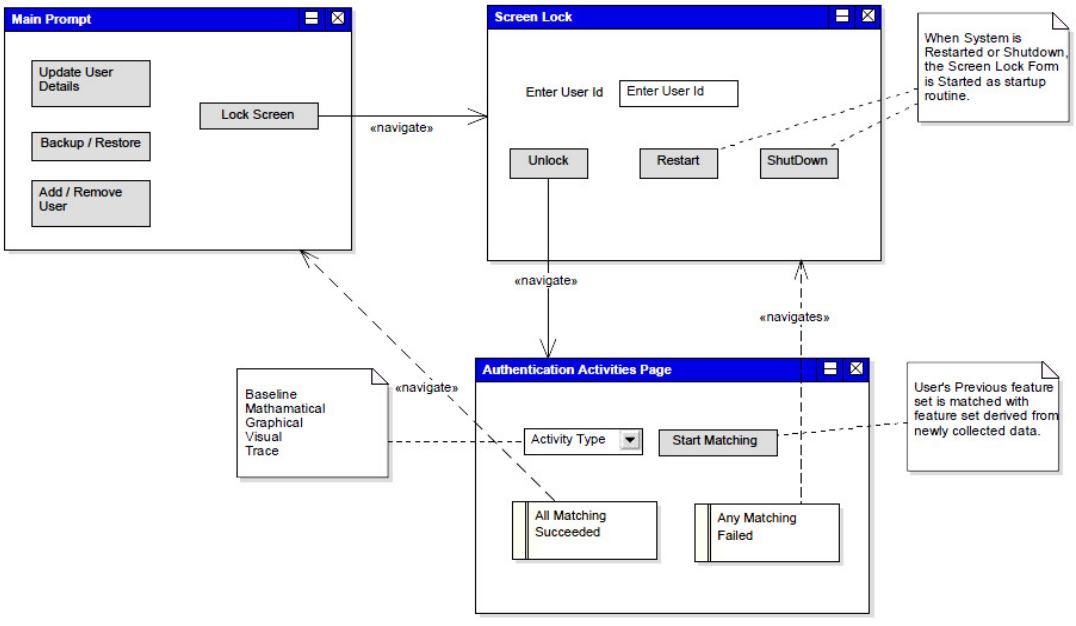


FIGURE 4: User Interface Diagram.

Step 2: We add a new user as there are no existing users initially. The training form opens wherein we train the system for authentication. The training is based on two activities, Meditation and Math activity. While the subject is performing these activities, the signals are recorded and stored.

Step 3: Once the training process is complete, the user returns to the main prompt form (Fig 5). The user can now lock the screen by clicking on the lock screen option. The login form appears wherein user name has to be specified for unlocking the screen (Fig 6). There are 3 available options, Unlock, Restart and Shutdown.

Step 4: When the unlock option is pressed by the user an authentication form appears. Two activities, for which the system has been trained earlier, must be performed for authentication, one after the other.

Step 5: If the authentication is successful then the main prompt form is displayed and the screen is successfully unlocked, else the authentication fails and the screen remains in the locked state.

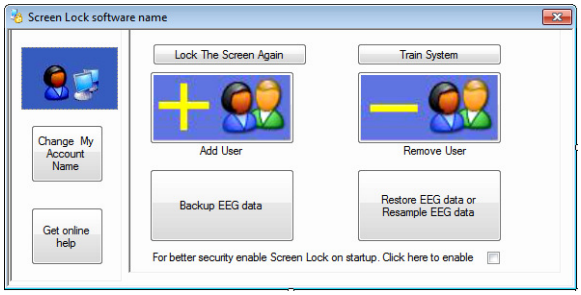


FIGURE 5: Main Prompt Window.

4. CONCLUSION

In this work, we investigate the use of brain activity for person authentication. It has been shown in previous studies that the brain-wave pattern of every individual is unique, and that the EEG can be used for biometric authentication. Person authentication aims to accept or reject a person claiming an identity. We perform EEG recording, feature extraction and matching of the feature vector with the stored feature vector, all in real time. This system seems to be the most reliable system of authentication as it is a type "*What I am"* system rather than the "What I Have" (Iris/Fingerprint scan) or "What I Know" (Password) variants of authentication system. Additionally, this system is designed without using any type of external stimulus. This work, however, needs more refinement such as,

i. Recording must be done in clinical conditions where there are no external interferences (noise free environment).

ii. Training the users to perform the various mental tasks with full concentration.

iii. Handling high dimensional data.

iv. Devising a more or less perfect matching algorithm that gives 0 FAE and 0 FRE.

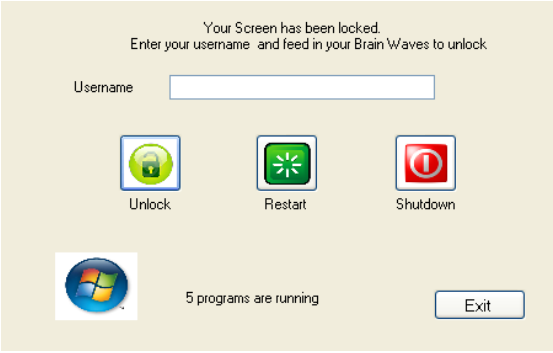


FIGURE 6: User login window

5. REFERENCES

[1] Palaniappan, R. (2008). "Two-stage biometric authentication method using thought activity brain waves." *International Journal of Neural Systems*, *18*(01), pp. 59-66.

[2] Bengio, S., Marcel, C., Marcel, S., & Mariéthoz, J. (2002). "Confidence measures for multimodal identity verification." *Information Fusion*, *3*(4), pp. 267-276.

[3] Abdullah, M. K., Subari, K. S., Loong, J. L. C., & Ahmad, N. N. (2010). "Analysis of the EEG Signal for a Practical Biometric System." *World Academy of Science, Engineering and Technology*, *68*, pp. 2067-2071.

[4] Gürkök, H., & Nijholt, A. (2012). "Brain–Computer Interfaces for Multimodal Interaction: A Survey and Principles." *International Journal of Human-Computer Interaction*, *28*(5), pp. 292-307.

[5] He, C., Lv, X., & Wang, Z. J. (2009, April). "Hashing the mAR coefficients from EEG data for person authentication." *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pp. 1445-1448.

[6] Marcel, S., & Millán, J. D. R. (2007). "Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation." *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *29*(4), pp. 743-752.

[7] Majumdar, K. (2011). "Human scalp EEG processing: Various soft computing approaches." *Applied Soft Computing*, 20*11*(8), pp. 4433-4447.

[8] Dieckmann, U., Plankensteiner, P., & Wagner, T. (1997). "Sesam: A biometric person identification system using sensor fusion." *Pattern recognition letters*, *18*(9), pp. 827-833.

[9] Xiong, N., & Svensson, P. (2002). "Multi-sensor management for information fusion: issues and approaches." *Information fusion*, *3*(2), pp. 163-186.

[10] Gao, J. B., & Harris, C. J. (2002). "Some remarks on Kalman filters for the multisensor fusion." *Information Fusion*, *3*(3), pp. 191-201.

[11] Tao, Q., & Veldhuis, R. (2009). "Threshold-optimized decision-level fusion and its application to biometrics." *Pattern Recognition*, *42*(5), pp. 823-836.

[12] Ross, A., & Jain, A. (2003). "Information fusion in biometrics." *Pattern recognition letters*, *24*(13), pp. 2115-2125.

[13] Lingaraju G M, Kusuma M, Vinay K, Rakshath K, Prakash S Y, Dharini R, "Person Authentication System Using Brain Waves as Biometric", *Conference on Evolutionary Trends in Information Technology*, Visvesvaraya Technological University, Belgaum, India, pp 47, 20-22nd May 2011 (CETIT2011).

[14] http://www.emotiv.com/eeg/features.php

[15] Poulos, M. Rangoussi, N. Alexandris, A. Evangelou, M. (2001). "On the use of EEG features towards person identification via neural networks." *Informatics for Health and Social Care*, *26*(1), pp. 35-48.

[16] del R Millan, J., Mouriño, J., Franzé, M., Cincotti, F., Varsta, M., Heikkonen, J., & Babiloni, F. (2002). "A local neural classifier for the recognition of EEG patterns associated to mental tasks." *Neural Networks, IEEE Transactions on*, *13*(3), pp. 678-686.

[17] Fatourechi, M., Bashashati, A., Ward, R. K., & Birch, G. E. (2007). "EMG and EOG artifacts in brain computer interface systems: A survey." *Clinical neurophysiology*, *118*(3), pp. 480-494.

[18] Hosni, S. M., Gadallah, M. E., Bahgat, S. F., & AbdelWahab, M. S. (2007, Nov). "Classification of EEG signals using different feature extraction techniques for mental-task BCI." In *Computer Engineering & Systems, 2007. ICCES'07. International Conference on* (pp. 220-226). IEEE.

[19] He, C., Lv, X., & Wang, Z. J. (2009, Apr). "Hashing the mAR coefficients from EEG data for person authentication." In *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on* (pp. 1445-1448).

[20] http://www.fon.hum.uva.nl/praat/manual/Principal_component_analysis.html

[21] Saa, J. F. D., & Gutierrez, M. S. (2010). "EEG Signal Classification Using Power Spectral Features and linear Discriminant Analysis: A Brain Computer Interface Application." LACCEI'2010, Innovation and Development for the Americas, Jun 1-4, 2010, Arequipa, Perú.

Multimodal Biometrics at Feature Level Fusion using Texture Features

Maya V. Karki *mayavkarki@msrit.edu*
Faculty, MSRIT, Dept. of E&C MSRIT
Bangalore-54, INDIA

Dr. S. Sethu Selvi *selvi_selvan@yahoo.com*
Faculty, MSRIT, Dept. of E&C MSRIT
Bangalore-54, INDIA

Abstract

In recent years, fusion of multiple biometric modalities for personal authentication has received considerable attention. This paper presents a feature level fusion algorithm based on texture features. The system combines fingerprint, face and off-line signature. Texture features are extracted from Curvelet transform. The Curvelet feature dimension is selected based on d-prime number. The increase in feature dimension is reduced by using template averaging, moment features and by Principal component analysis (PCA). The algorithm is tested on in-house multimodal database comprising of 3000 samples and Chimeric databases. Identification performance of the system is evaluated using SVM classifier. A maximum GAR of 97.15% is achieved with Curvelet-PCA features.

Keywords: Multimodal Biometrics, Feature Level, Curvelet Transform**,** Template Averaging, PCA Features and SVM Classifier.

1. INTRODUCTION

Personal authentication systems built upon only one of the biometric traits are not fulfilling the requirements of demanding applications in terms of universality, uniqueness, permanence, collectability, performance, acceptability and circumvention. This has motivated the current interest in multimodal biometrics [1] in which multiple biometric traits are simultaneously used in order to make an identification decision. Depending on the number of traits, sensors and feature sets used, a variety of scenarios are possible in a multimodal biometric system. They include single biometric with multiple sensors, multiple biometric traits, single biometric with multi-instances, single biometric with multiple representations and single biometric with multiple matchers. Among all these scenarios, system with multiple biometric traits is gaining importance and this method itself is known as multimodal biometric system. Based on the type of information available in a certain module, different levels of fusion are defined [2]. Levels of fusion are broadly classified into two categories: fusion before matching also called as pre-classification which includes sensor level and feature level. Fusion after matching also called as post classification which includes match score level and decision level. Amongst these, fusion at feature level is gaining much research interest.

Most of the existing multimodal systems are based on either score level or decision level fusion [3]. Match score is a measure of the similarity between the input and template biometric feature vector. In match score level, scores are generated by multiple classifiers pertaining to different biometric traits and combined [4]**.** In order to map score of different classifiers into a single domain, where they possess a common meaning in terms of biometric performance, normalization technique is applied to the output of classifier before score fusion. Gupta [5] developed a multimodal system based on fingerprint, face, iris and signature with score level

fusion. In all these systems texture features are extracted and score level and decision level fusion are compared using SVM classifier. The most promising recent research is certainly the information fusion at the matching score level invoking user specific weights and threshold levels. Though a few multimodal systems developed are considered to be very accurate, still they are not validated since, systems are tested on a medium size database.

Fusion at feature level involves integration of feature sets corresponding to multiple biometric traits. Feature set contains rich information about biometric data than the match score or final decision. Therefore integration at this level is expected to give improved recognition performance. Due to the constraints of feature level fusion, very few researchers have studied integration at feature level. Chetty [6] combined face and voice using visual and acoustic features with artificial neural network as a recognizer and obtained an Equal Error Rate (EER) of 2.5%. Nandakumar [7] concatenated fingerprint and iris code at feature level using fuzzy vault classifier and showed that uncorrelated features when combined gives best possible EER. Ferrer [8] proposed fusion of features extracted from hand geometry, palmprint and fingerprint. It is imperative that an appropriate feature selection scheme is used when combining information at the feature level.

This paper proposes a multimodal identification system that combines fingerprint, face and signature biometric traits. These three traits are considered due to their wide acceptance by users and also the data acquisition cost involved in these three traits are much less compared to other biometrics. Texture features are extracted from each modality independently and fusion at feature level is performed. Texture features are extracted from Curvelet Transform. Section 2 describes the proposed multimodal biometric Identification system based on feature level fusion. Section 3 describes database collection protocol and pre-processing. Section 4 describes feature extraction and dimension reduction techniques. Section 5 summarizes experimental results and section 6 gives comparisons with similar work. Section 7 concludes proposed system.

2. PROPOSED MULTIMODAL IDENTIFICATION SYSTEM

The schematic of the multimodal system at feature level fusion is shown in Figure 1. The multimodal system has two phases of operation: enrolment and identification. The enrolment module registers a person and then the three biometric traits are acquired and representation of these three traits are stored in a database. The proposed system is designed to operate in parallel mode. Fingerprint, face and signature of a person are acquired separately in data acquisition module. Required pre-processing techniques are applied on every biometric trait and features are extracted simultaneously. Features from all three biometric traits are concatenated and a feature vector is formed and stored as template in a database. Single matcher is used to evaluate the performance. SVM classifier is used for matching. During authentication, feature vector extracted from the test person is compared with the template stored in the database. Matching is performed using a recognizer which compares query feature vector with the template in the database and generates a match score. To prevent impostor from being identified, the match score from matcher is compared with a predefined threshold in the decision module. This module makes a decision as either person under test is recognized or not recognized.

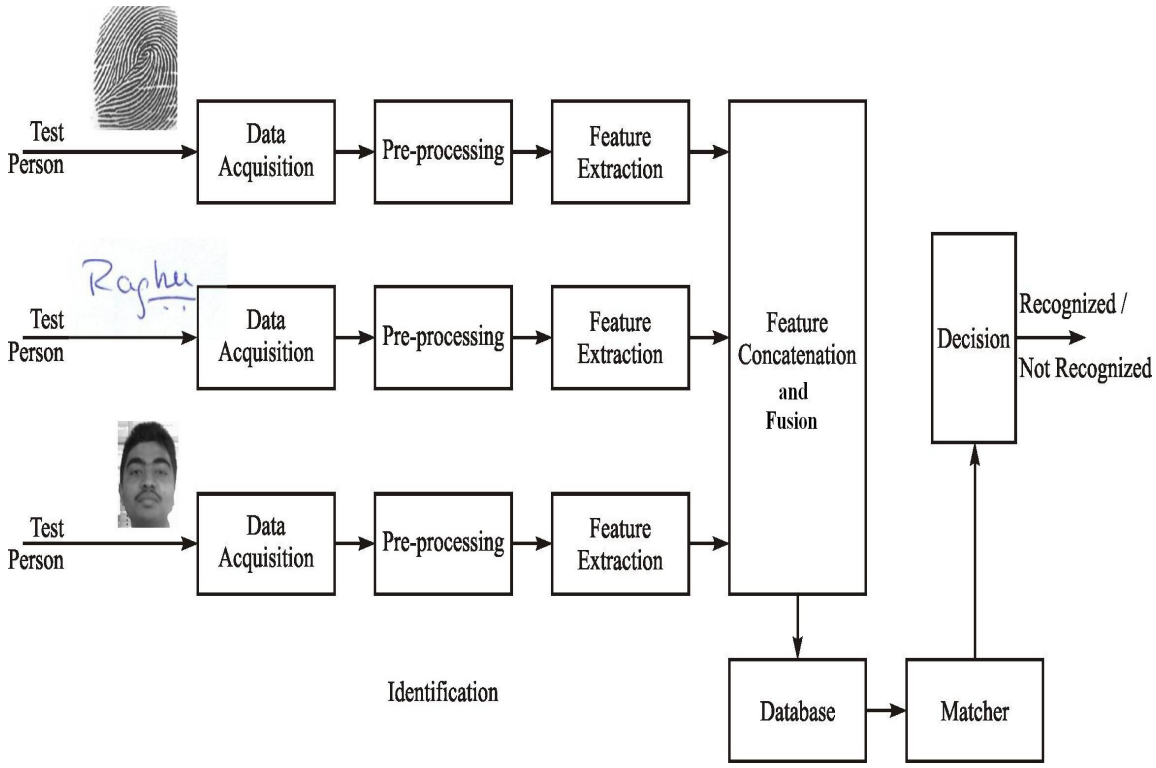


FIGURE 1: Block Diagram of Proposed Multimodal Biometric System Based on Feature Level Fusion.

2. DATABASE COLLECTION and PRE-PROCESSING

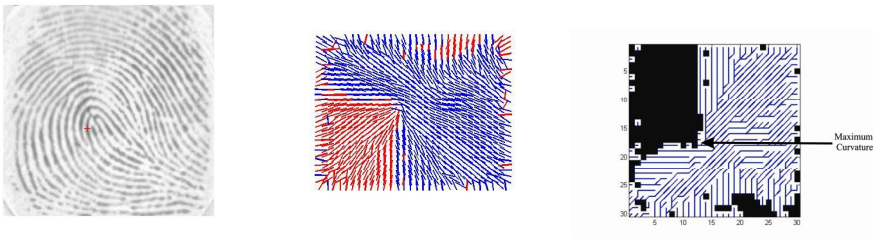
A multimodal database including fingerprint, signature and face samples are essential to test the performance of the proposed system. Since there is no standard database freely available to meet the requirement of the proposed algorithm, ECMSRIT multimodal database and Chimeric databases have been formed. ECMSRIT database is collected from fingerprint, off-line signature and face samples of 100 users. Collection of these unimodal traits are described below:

Nitgen fingerprint scanner is used to collect fingerprints. It is an optical sensor with ultra-precise 500dpi resolution. To locate centre point of fingerprint, it is divided into non-overlapping blocks. Gradient in x and y direction at each pixel in a block is obtained. A 2D Gaussian filter is applied to smooth the gradient. A slope perpendicular to direction of gradient in each block is computed. Blocks with slope values ranging from 0 to pi/2 are considered. In each block a path is traced down until a slope that is not ranging from 0 to pi/2 and that path is marked. Block with highest number of marks gives slope in the negative y direction. This provides the centre point of fingerprint. Region of interest around the centre point is cropped and normalized in size to 128 *128 pixels. Figure 2 represents centre point detection and cropping of fingerprint. Figure 2 (a) represents scanned fingerprint, (b) shows orientation of fingerprint, (c) represents maximum curvature points, (d) shows centre point and (e) shows cropped fingerprint and (f) shows third level LL subband of Curvelet transformed fingerprint.

Still face images are collected using digital camera LifeCam Nx-6000 with a 2.0 mega pixels sensor. The 2D colour face image is converted to a gray scale image. Canny edge detection mask with suitable threshold value is applied on image with a uniform background to extract outer curvature of the face. From this only foreground face image of size 128 * 128 is cropped Figure 3 (a) shows edge detection and (b) represents cropped face and (d) represents third level LL subband of Curvlet transformed face.

The signatures were taken on a A-4 size white paper. These were scanned using an 8-bit, 300 dpi resolution scanner. The scanned signatures were cut out from the scanned page in their original orientation and size using an image editor. The scanned signature is binarized. Since the signature consists of black pixels on a white background, the image is then complimented to make it a white signature on a black background. When a signature is scanned, the image obtained may contain some noise components like the background noise pixels and these noise pixels are removed by employing median filter. To avoid inter-personal and intra-personal size variations of signatures, size is normalized to 128 * 256. Figure 4 (a) shows input signature, (b) noise removed, (c) complemented and (d) represents normalized signature sample.

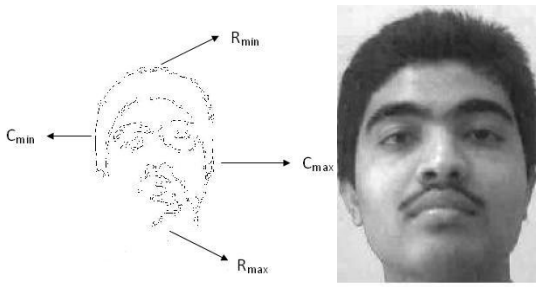
The size of MSRIT database is 10 x 3 x 100= 3000. Chimeric database-I is formed by FVC2002-DB3 fingerprint, ECMSRIT signature and ORL face databases. As ORL face database has only 40 users, Chimeric database-I is formed by considering only 40 users from fingerprint and signature databases. Chimeric database-I consists of 8 samples of each person for each trait with total of 8 x 3 x 40 = 960. Chimeric database-II is formed by FVC2004-DB3 fingerprint, CEDAR signature and Faces-94 face databases. As the CEDAR signature database has only 55 users, Chimeric database-II is formed by considering only 55 users from fingerprint and face databases. FVC2004-DB3 has only 8 samples per user and Chimeric database-II consists of 8 samples of each person for each modality with total of 8 x 3 x 55=1320 samples.



a) Input fingerprint (b) Orientation (c) Maximum curvature points

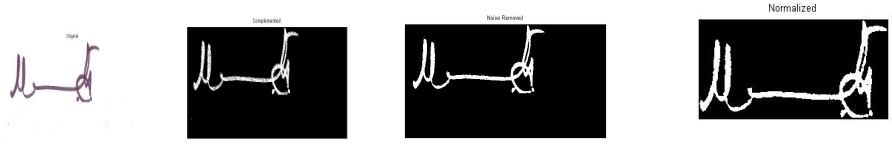
(d)Fingerprint with centre (e) Cropped fingerprint

FIGURE 2: Result of Pre-Processing of Fingerprint.



(a) Edge Detection. (b) Cropped Face.

FIGURE 3: Result of Pre-Processing of Face.



a) Input Signature (b) Noise Removed (c) Complemented (d) Normalised

FIGURE 4: Result of Pre-Processing of Signature.

4. FEATURE EXTRACTION

4.1 Extraction of Texture Features using Curvelet Transform

Texture features are extracted by applying Curvelet transform on each trait. Curvelet transform [10,11,12] is based on multi-scale ridgelet transform [12] combined with spatial bandpass filtering operation at different scales. It is better for representing point discontinuities. Figure 5 shows the flow graph of Curvelet transform. The transform involves following steps: (1) The subbands of input trait is obtained using DB4 wavelet transform. (2) The 2D fast Fourier transform of the LL subband is obtained. (3) Using interpolation scheme, the samples of the Fourier transform obtained on the square lattice is substituted with sampled values on a polar lattice. (4) Inverse fast Fourier transform is computed on each radial line. (5) 1D Wavelet transform is computed at each radial line using DB4 filter and approximate coefficients are used as features.

Steps 2 through 5 form Ridgelet transform [12] and Steps 2 to 4 represent finite Radon transform[13] for digital data. Figure 6 (a), (b) and (c) show third level LL subband of Curvelet transformed fingerprint, face and signature respectively. For example, consider a normalized signature of size 128 x 256. Following the steps described above, third level Curvelet transformed LL subband coefficients of size 25 x 20 will give feature dimension of 500.

Curvelet feature dimension is decided based on d-prime number (d'). Performance of the biometric system has been predicted by calculating d' value [3]. It measures separation between the means of the genuine and impostor probability distributions in standard deviation unit. To evaluate d', genuine and impostor match scores are calculated. A match score is found to be genuine if it is a result of matching two samples of the same user and is known as impostor score if two samples of different users are compared. During training period all samples in the database are considered to find genuine and impostor score. Let P be the number of persons enrolled in

the system and let S be the number of samples of the trait obtained from each person, then number of genuine scores G_{score} and number of impostor scores I_{score} are given by

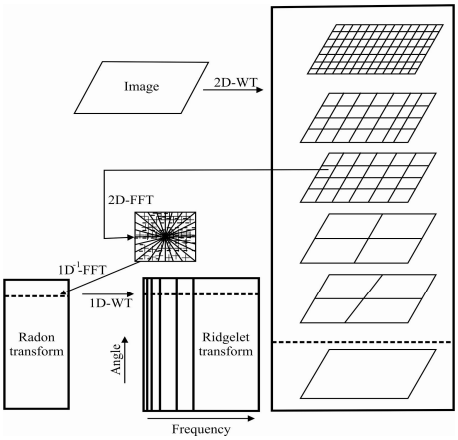


FIGURE 5: Flow Graph of Curvelet Transform.

$$G_{score} = P \times (S-1) \times \frac{S}{2} \quad (1)$$

$$I_{score} = P \times S^2 \times \frac{(P-1)}{2} \quad (2)$$

From these genuine and impostor scores mean and standard deviation are calculated to evaluate d'. The d' value is given by

$$d' = \sqrt{2}\,\frac{\mu_{genuine} - \mu_{impostor}}{\sqrt{\sigma^2_{genuine} + \sigma^2_{impostor}}} \quad (3)$$

Where μ and σ are the mean and standard deviation of genuine and impostor scores. For each trait, different dimension of Curvelet features are evaluated and corresponding d' value has been calculated. Figure 7 shows the variation of d' value for different feature dimensions. From the graph it is observed that as feature dimension increases, d' value increases and higher the value of d' better is performance. The d' value remains constant for a feature dimension of 500 and above. Hence, for each trait a maximum feature dimension of 504 has been considered to evaluate recognition performance of the system. In feature level fusion the features from each trait are concatenated. With concatenation feature dimension increases and to reduce the dimension few reduction techniques [14] are adapted.

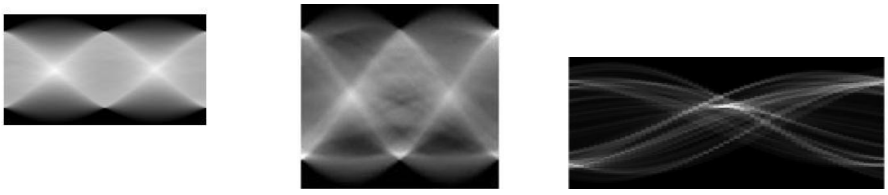


FIGURE 6: Third Level LL Subband of (a) Fingerprint (b) Face and (c) Signature.

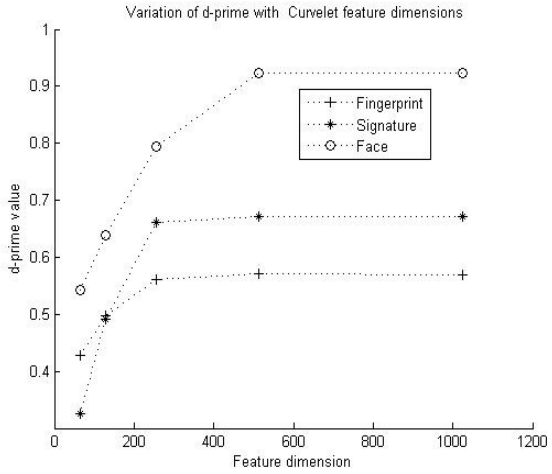


FIGURE 7: Variation of d' Value for Different Feature Dimension.

4.2 Dimension Reduction Techniques

Let FP, FS and FF be three feature vectors extracted by applying Curvelet transform on fingerprint, signature and face respectively. Let fp, fs and ff be the dimension of each trait. Feature vectors of three traits are represented by

$$FP = [P_1, P_2 ..... fp]$$

$$FS = [S_1, S_2 ....., fs]$$

$$FF = [F_1, F_2 ....., Ff] \quad (4)$$

All three feature vectors are concatenated to form a new feature vector F_c where $F_c = FP + FS + FF$

The dimension of F_c is equal to fp+fs+ff. In the proposed algorithm approximate coefficient features of dimension 504 is extracted from each of the trait and concatenated, resulting in dimension 504+504+504= 1512. F_c is stored as new template in the database for matching.

(1) Template Averaging: The concatenated feature vector shows increase in size but homogeneous in nature. Therefore the size of concatenated feature vector F_c is reduced by

applying averaging and this method is known as template averaging [3] and this method is very simple. Average of feature vector calculated from three traits is

$$F_a = \frac{FP + FS + FF}{3} \qquad (5)$$

(2) PCA Features: In this method, to reduce the dimension of concatenated feature vector, Principal Component Analysis (PCA) [15] is applied to F_c. PCA [16] transforms a number of correlated variables into number of uncorrelated variables referred as principal components. PCA reduces dimension without loss of information and preserves the global structure. Two dimensional PCA (2DPCA) [14] has been applied on the concatenated features from three traits. Algorithmic steps involved in applying 2DPCA are (1) Subband coefficients from Curvelet transform are extracted from each trait and concatenated. Let A be the concatenated matrix. Covariance of matrix of A is calculated. (2) Eigen values and eigen vectors of covariance matrix are calculated. Eigen vector with highest eigen value is called as principal component of the matrix A. Choosing first v eigen values from covariance matrix A, a transformed matrix B is obtained as B=A x P where P=[P_1,P_2 …Pv] is the projection matrix whose columns are the eigen vectors of covariance matrix in the decreasing order of the eigen values. B is the required feature matrix which is stored as PCA feature vector F_{pca} in the database for matching. By applying Curvelet transform for fingerprint, face and signature, subband matrix of size 18x28, 18x28 and 20x25 are extracted and concatenated to form a matrix A of size 60x 25. 2DPCA is applied on A and by considering first eight eigen values, the transformed matrix 60x8 gives a feature dimension of 480.

(3) PCA Features without Fusion: To reduce the dimension at feature level fusion, 2DPCA is applied on the subbands of Curvelet transform of each trait independently. The subband PCA features are concatenated to form a feature vector F_p and stored as templates in the database. 2DPCA is applied on Curvelet subband obtained from each trait independently and selecting nine largest eigen vectors from fingerprint, face and eight eigen vectors from signature, PCA feature of dimension 162, 162 and 160 is obtained and concatenated to form feature dimension of 484.

(4) Statistical Moment Features without Fusion: Moment features are extracted from subbands of Curvelet transform [18] from each trait and calculated as described below

Mean of each subband is calculated. Let μ_k be the mean value of k^{th} subband. Second order moment or variance of each subband σ_k is calculated using

$$\sigma_k = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(W_k(i,j) - \mu_k \right)^2 \qquad (6)$$

Where W_k indicates the subband coefficients of k^{th} ban and MxN be the size of subband.

Third order moment is calculated as

$$\mu_{3k} = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(W_k(i,j) - \mu_k \right)^3 \qquad (7)$$

Fourth order moment is calculated as

$$\mu_{4k} = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \left(W_k(i,j) - \mu_k \right)^4 \qquad (8)$$

Energy of each subband is calculated as

$$E_k = \sum_{i=1}^{M}\sum_{j=1}^{N} W_k^2(i,j) \quad (9)$$

The resulting feature vector of k^{th} subband is given by

$F_{mk} = [\ \mu_k\ \sigma_k\ \mu_{3k}\ \mu_{4k}\ E_k\ ]$

These moment features from each trait are concatenated to form a new feature vector F_m and stored in a database. First, second and third level Curvelet decompositions are applied on each trait and results into 12 subbands. From these 12 subbands moment features of size 60 are calculated and when concatenated from three traits give a feature dimension of 180.

(5) Feature Concatenation by extracting Significant Coefficients: Curvelet subband coefficients are extracted from each trait and sorted. Significant coefficients from each trait are selected and concatenated to form a feature vector F_r. Dimension of F_r is made comparable to feature dimension of unimodal system. For example, by applying Curvelet transform on each trait 504 subband coefficient features are sorted and only first 168 features from each trait are concatenated to form a feature vector of dimension 168+168+168= 504.

5. EXPERIMENTAL RESULTS

Performance of the proposed algorithm is tested for identification mode using SVM classifier. In SVM classifier method [19, 20], each person in the database has an associated SVM. The test samples are assigned the label of the person whose SVM gives the largest positive output. SVM classifier with a polynomial kernel of order 2 is selected. Penalty parameter C is tuned from 2 to 10 to get better results. In this experiment, C is set to a value 2.

Samples in each of the databases are split into training and test set. Training and test samples are selected in different ratios starting from 1:9,2:8,3:7 ....,9:1 and corresponding recognition rate for five trials have been calculated using Euclidean distance measure. The average recognition rate is calculated and result is compared with the results obtained from different sets of train and test ratios and the ratio which gives maximum recognition rate is considered for performance evaluation. In this experiment train to test ratio of 6:4 is considered and each database is randomly split 30 times at each time performance of the system is evaluated and average of these 30 times result is considered as final result. Genuine and impostor scores are calculated for the six different feature vectors. Figure 7 shows histogram plots evaluated on ECMSRIT database for six feature sets.

From figure8 it is seen that separation of genuine and impostor scores are more in F_{pca} and F_a features compared to other features. Based on these distribution curves, the threshold is varied to calculate FAR and FRR from which EER has been calculated for each of the feature vector. Figure 9 shows threshold vs FAR and threshold vs FRR for all feature vectors evaluated. Figure 9 indicates FAR and FRR varies for each algorithm and each feature vector, resulting to variation in EER. Table 1 indicates a minimum EER of 5.32% is obtained for F_{pca} and a maximum EER of 22.35% is obtained for F_r features.

Features	Feature Dimension	Optimal Threshold	FAR (%)	FRR (%)	EER (%)
F_c	1512	0.46	19.85	18.54	19.31
F_a	504	0.45	12.07	11.78	12.00
F_{pca}	480	0.38	5.34	4.35	5.32
F_p	484	0.31	16.46	15.53	15.33
F_m	180	0.14	20.54	17.78	20.54
F_r	504	0.35	25.25	22.34	22.35

TABLE 1: EER (%) for Different Features.

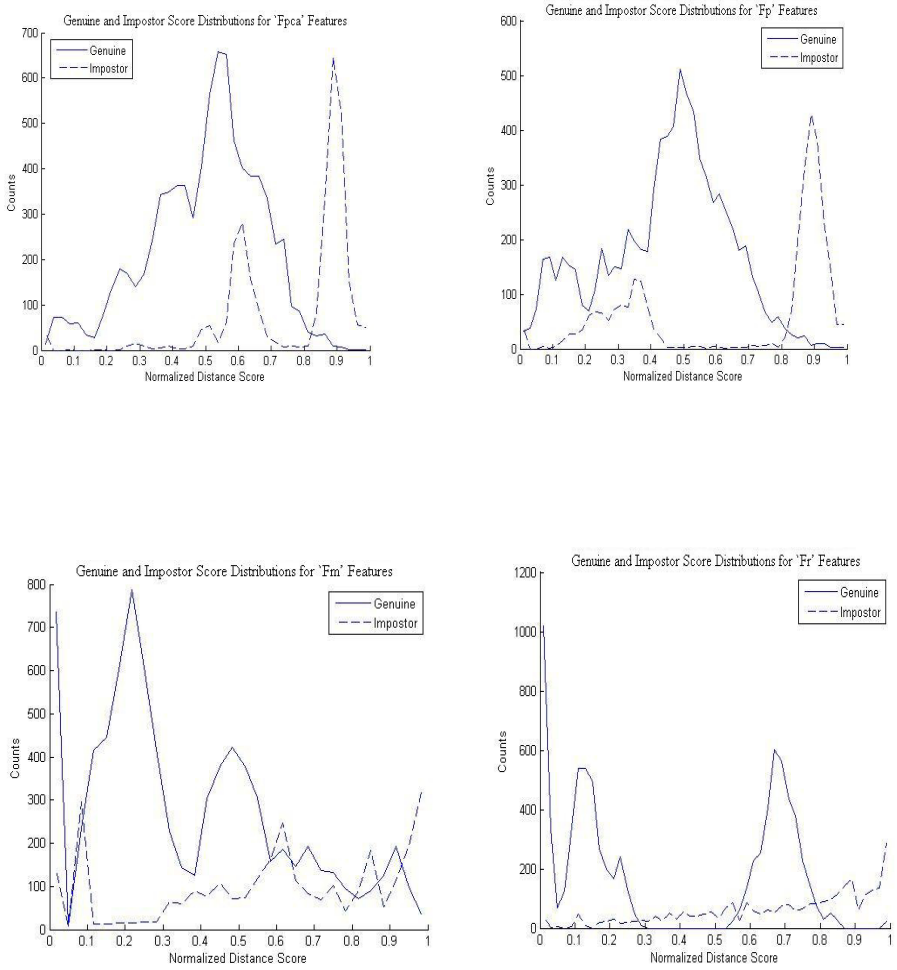


FIGURE 8: Histogram Plots for Genuine and Impostor Scores.

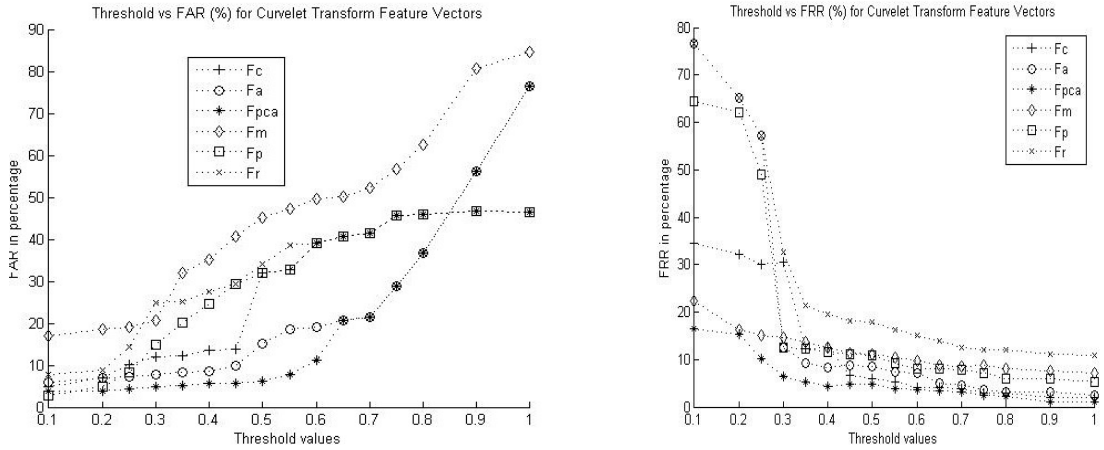


FIGURE 9: FAR and FRR vs Threshold for Different Feature Sets.

Biometric Traits	Feature Vector	Feature Dimension	GAR (%)
Fingerprint+ Signature Fingerprint +Face Face +Signature	F_c	504+504=1008	94.52 94.36 93.14
Fingerprint+ Signature Fingerprint+ Face Face+ Signature	F_a	504	93.52 95.26 92.24
Fingerprint+ Signature Fingerprint+ Face Face+ Signature	F_{pca}	480	93.82 96.64 96.16
Fingerprint+ Signature Fingerprint+ Face Face+ Signature	F_p	484	94.08 94.86 93.04
Fingerprint+ Signature Fingerprint+ Face Face+ Signature	F_m	120	89.05 90.05 88.26
Fingerprint+ Signature Fingerprint+ Face Face+ Signature	F_r	504	89.65 90.45 89.26

TABLE 2: Performance of Identification based on Curvelet Feature Vectors by Combining Two Traits in ECMSRIT Multimodal Database.

Multimodal Database	Feature Vector	Feature Dimension	GAR(%)
ECMSRIT Chimeric Database-I Chimeric Database-II	F_c	504+504+504=1512	96.82 96.04 96.84
ECMSRIT Chimeric Database-I Chimeric Database-II	F_a	504	96.92 95.22 95.02
ECMSRIT Chimeric Database-I Chimeric Database-II	F_{pca}		97.15 96.32 96.14
ECMSRIT Chimeric Database-I Chimeric Database-II	F_p		96.08 94.93 94.82
ECMSRIT Chimeric Database-I Chimeric Database-II	F_m		92.67 92.35 92.86
ECMSRIT Chimeric Database-I Chimeric Database-II	F_r		92.35 91.25 91.89

TABLE 3: Performance of Identification based on Curvelet Feature Vectors by Combining Three Traits.

Table 2 shows the performance of the system at feature level fusion considering two traits at a time. Results show that GAR obtained from fingerprint and face is more compared to other two combinations and maximum GAR of 96.64% is obtained for F_{pca} features. Table 3 shows the performance at feature level fusion from all three biometric traits and a maximum GAR of 97.15% is obtained for F_{pca} features. Results show that GAR is better in ECMSRIT database compared to Chimeric databases. This is because in Chimeric databases the three biometric traits are not from the same person. The performance of the proposed algorithm is better for correlated traits compared to non-correlated traits. The test samples are rotated in steps of 2^o to verify rotation invariance for F_c, F_{pca} and F_a. Feature sets as these feature sets give better GAR compared to other three feature sets. The results show that the GAR for rotated samples also remains almost same and confirms that Curvelet transform is rotation invariant.

Feature Vector	Feature Dimension	Rotation in Degrees					
		0^o	2^o	4^o	6^o	8^o	10^o
GAR (%) for ECMSRIT Multimodal Database							
F_c	1512	96.82	96.45	95.20	95.04	94.86	94.32
F_a	504	95.92	95.22	94.74	94.22	93.75	93.12
F_{pca}	480	96.78	96.22	95.74	94.22	93.75	93.12
GAR (%) for Chimeric Database-I							
F_c	1512	96.04	96.04	95.86	94.75	93.25	92.89
F_a	504	95.22	95.22	94.74	94.22	93.75	93.12
F_{pca}	480	96.32	95.02	94.78	93.21	92.56	92.12
GAR (%) for Chimeric Database-I I							
F_c	1512	96.84	95.05	94.54	93.84	92.14	91.85
F_a	504	95.02	94.89	94.74	94.22	93.75	93.12
F_{pca}	480	96.14	95.42	94.02	93.65	92.28	91.64

TABLE 4: Recognition Results by Applying Rotation to Test Samples for three Feature Sets.

Figures 9 shows the comparison between GAR obtained from unimodal identification system with subband coefficient features and multimodal identification system with F_{pca} features. F_{pca} gives maximum GAR and its dimension remains same as that of unimodal traits. When fingerprint is combined with face and signature an improvement in GAR of 9.09% and when signature is combined with face and fingerprint improvement has been 12.67%. Chart also show that improvement in GAR is less when performance of two and three traits are compared.

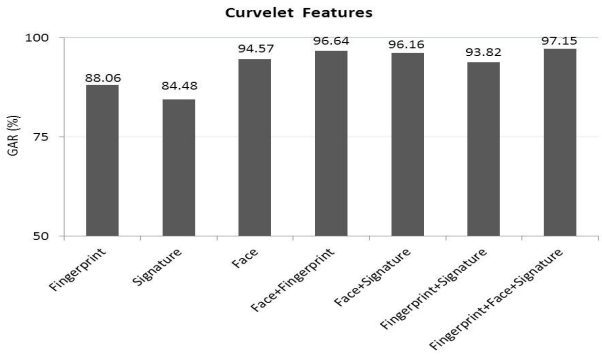


FIGURE 10: Comparison between GAR Obtained from Unimodal and Feature Level Identification for Curvelet Features.

6 COMPARISONS WITH SIMILAR WORK

Currently, as there are few multimodal system combining fingerprint, signature and face with curvelet features at feature level unimodal recognition system is considered for comparison. Table 5 gives comparison of fingerprint, face and signature recognition system based on Curvelet features. In [21] author applied Curvelet transform on a fingerprint of size 64 x 64 which was divided into four blocks. Each block is divided into 8 angular directions and standard deviation from each direction was concatenated to form feature vector. Proposed Curvelet algorithm with moment feature of dimension 40 is used for comparison and GAR obtained for 15 users is higher. In [22], author applied sixth level Curvelt transform decomposition and used 160 subband coefficients as features. Performance is evaluated using SVM classifier. The proposed algorithm with 120 subband coefficients as features has similar performance. In [23], author applied Curvelt transform on face at different scales subband coefficients are used as features. Few dimension reduction techniques are applied and performance is evaluated. GAR obtained

with the subband feature dimension of 1258 is similar to the GAR obtained from proposed algorithm with the feature dimension of 504.

Unimodal Recognition System	Author	Database used	No. of users	Feature Dimension	Classifier	Performance
Fingerprint Verification	A.Mujumadar[21]	FVC2004-DB1	15	32	Fuzzy-KNN	GAR=91.7%
	Proposed Algorithm	FVC2004-DB3	15	40	Euclidean	GAR=95.02%
Signature Verification	M.Fakhlai [22]	Own	39	160	SVM	GAR=89.87%
	Proposed Algorithm	CEDAR	55	120	Euclidean	GAR=88.87%
Face Recognition	Tanaya G. [23]	ORL	40	1258	Euclidean	GAR=94.54%
	Proposed Algorithm	ORL	40	504	SVM	GAR=95.04%

Table 5: Comparison of Unimodal Recognition System based on Curvelet Features.

7. CONCLUSION

The proposed multimodal system comprises of fingerprint, off-line signature and face traits. Performance of the system is evaluated based on Curvelet transform features with SVM classifier. The algorithm is tested by combining two traits at a time and three traits together. The increase in dimension at feature level fusion is reduced by using template averaging, PCA and statistical moment features. Six different feature vectors from these algorithms have been tested at feature level fusion. F_{pca} and F_a features are obtained from concatenated feature vector while F_p, F_m and F_r feature vectors are obtained without any fusion. The dimension of reduced feature vectors are comparable with those of unimodal traits. The feature dimensions of unimodal traits are decided based on d' values. Fusion algorithm is tested on in-house created ECMSRIT multimodal and Chimeric databases. Though the size of databases are small, the performance obtained from these databases are low compared to that of ECMSRIT database. Results indicate that the proposed algorithm performs better on correlated data than uncorrelated data. From simulation results it can be summarized that when three traits are combined performance of the system increases compared to either unimodal system or by combining two traits.

8. ACKNOWLEDGEMENTS

This work was supported in part by Research Grants from AICTE, New Delhi under Research Promotion Scheme, grant no. 8020/RID/BOR/RPS-42/2005-2006.

9. REFERENCES

[1] Anil K. Jain and Arun Ross. Multimodal Biometrics an Overview. Communications of ACM, pages 1221{1224, September 2004.

[2] K. Nanadakumar. Multibiometric System: Fusion Strategies and Tem plate Security. PhD thesis, MSU, 2008.

[3] Arun Ross, Nandakumar, and A.K.Jain. Handbook of Multibiometrics. Springer Verilag New York, 1 edition, 2004.

[4] R. Snelick, U. Uludag, A. Mink, M. Indovina, and A. K. Jain. Large Scale Evaluation of Multimodal Biometric Authentication using State of the art Systems. IEEE Transactions On Pattern
[5] Phalguni Gupta, Ajita Rattani, aHunny Mehrotra, and Anil Kumar Kaushik. Multimodal Biometrics System for Efficient Human Recognition. In Proceedings of SPIE, 2006.Analysis and Machine Intelligence, 27(3):450{455, 2005.

[6] Girija Chetty and Wagner M. Investigating Feature Level Fusion for Checking Liveness in Face-Voice Authentication. In Proceedings of the Eighth International Symposium on Signal Processing and Applications, ISSPA-2005, pages 66{69, August 2005.

[7] K. Nandakumar and A. K. Jain. Multibiometric Template Security using Fuzzy Vault. In Proceedings of IEEE Second International Conference on Biometrics: Theory, Applications and Systems, September 2008.

[8] Ferrer, Miguel A.and Travieso, Carlos M.and Alonso, and Jesus B. Multimodal Biometric System based on Hand geometry and Palmprint Texture. In Proceedings 40th Annual IEEE International Carnahan Confrence on Security, pages 92{97,2006.

[9] M. Husken, S M. Brauckmann, K. Okada Gehlen, and C. V. Malsburg. Evaluation of Implicit 3D Modeling for Pose-invariant Face Recognition, 2004.

[10] D..Donoho and M.R.Duncun. Digital Curvelet Transform: Strategy, Implementation and Experiments. Technical report, Stanford University, 1999.

[11]Guillaume, Joutel, Eglin, and Bres Emptoz. Curvelet based Feature Extraction of Handwritten Shapes for Ancient Manuscripts Classi_cation. In Procedings Of SPIE Electronic Imaging, volume 6500, pages 1-12, 2007.

[12]Jianwei Ma and Gerlind Plonka. A Review of Curvelets and Recent Applications. IEEE Signal Processing Magazine, 27(2):118{133, 2010.

[13]Miroslaw Miciak. Radon Transformation and Principal Component Analysis, Method Applied in Postal Address Recognition Task. International Journal of Computer Science and Applications, 7(3):33{34, 2010.

[14] Cheng Lu and Liu. Multimodal Biometrics Recognition by Dimensionality Reduction Methods. In Second International Symposium on Electronics Commerce and Security, pages 113{116, 2009.

[15] M. Turk and A. Penteland. Eigenfaces for Recognition. Cognitive Neuroscience, 3(1):71{86, 1991.

[16] G.C.Feng, P.C.Yuen, and D.Q.Dai. Human Face Recognition using Wavelet Subband. Journal of Electronic Imaging, 2(2):226{233, 2000.

[18] F.Murtagh and J.L.Starck. Wavelet and Curvelet Moments for Image Classification: Application toAggregate Mixture Grading. Journal on Computer Vision and Pattern Recognition, 29:1557{1564, 2008.

[19] V. Vapnik. The Nature of Statistical Learning Theory. Springer-Verlag, New York, NY, 1995.

[20] Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology, vol 2 Issue 3, article no. 7:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.

[21] A.Mujumadar and R.K.Ward. Fingerprint Recognition with Curvelet Features and Fuzzy KNN Classi_er. In Proceedings of Signal and Image Processing, 2008.

[22] M.Fakhlai and H.Pourreza. O_-line Signature Recognition based on Wavelet, Curvelet and Contourlet transforms. In International Conference on Document Analysis and Recognition (ICDAR), pages 734{738, September 2007.

[23] Tanaya Guha, Q.M. Jonathan Wub and Yuan Yuan" Curvelet based face recognition via dimension reduction" Signal Processing Volume 89, Issue 12, page no. 2345-2353

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Biometric and Bioinformatics (IJBB)* brings together both of these aspects of biology and creates a platform for exploration and progress of these, relatively new disciplines by facilitating the exchange of information in the fields of computational molecular biology and post-genome bioinformatics and the role of statistics and mathematics in the biological sciences. Bioinformatics and Biometrics are expected to have a substantial impact on the scientific, engineering and economic development of the world. Together they are a comprehensive application of mathematics, statistics, science and computer science with an aim to understand living systems.

We invite specialists, researchers and scientists from the fields of biology, computer science, mathematics, statistics, physics and such related sciences to share their understanding and contributions towards scientific applications that set scientific or policy objectives, motivate method development and demonstrate the operation of new methods in the fields of Biometrics and Bioinformatics.

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 IJBB.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Started with Volume 7, 2013, IJBB appears with more focused issues related to biometrics and bioinformatics studies. Besides normal publications, IJBB intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

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

LIST OF TOPICS

The realm of International Journal of Biometrics and Bioinformatics (IJBB) extends, but not limited, to the following:

- Bio-grid
- Bioinformatic databases
- Biomedical image processing (registration)
- Biomedical modelling and computer simulation
- Computational intelligence
- Computational structural biology
- DNA assembly, clustering, and mapping
- Fuzzy logic
- Gene identification and annotation
- Hidden Markov models
- Molecular evolution and phylogeny
- Molecular sequence analysis
- Bio-ontology and data mining
- Biomedical image processing (fusion)
- Biomedical image processing (segmentation)
- Computational genomics
- Computational proteomics
- Data visualisation
- E-health
- Gene expression and microarrays
- Genetic algorithms
- High performance computing
- Molecular modelling and simulation
- Neural networks

CALL FOR PAPERS

Volume: 7 - **Issue: 2**

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