Enhanced Feature Extraction System using ICA, GA and GFCC for Ear, Iris and Speech

Harpreet Kaur[1], Ranjit Singh[2]
[1]*Assistant Professor (Computer Science & Engg),* [2]*Associate Professor (Applied Sciences)*
Chandigarh Group of Colleges, Jhanjeri, Mohali, Punjab.

Abstract **-** Biometric technologies are becoming the foundation of an extensive array of highly secure identification and person verification solution. Multimodal biometric identification systems aim to fuse two or more traits. In this proposed work, physiological biometrics from iris and ear are combined with behavioral biometrics from speech to achieve robust user authentication on feature level fusion. Gammatone Frequency Cepstral Coefficient (GFCC) is used for speaker recognition. Independent Component Analysis (ICA) and Enhanced Genetic Algorithm (GA) for iris, ear feature extraction and feature reduction. Results of the three modalities were fused by using mean Mahalanobis distance. Finally the feature vector is matched with stored template using Hamming distance and provides optimal False Acceptance Rate (FAR) and False Rejection Rate (FRR), thus improving system accuracy.

Keywords **-** Multi-modal biometric, Morphological operation, Independent Component Analysis, Genetic Algorithm, Mahalanobis distances, hamming distance. Feature level fusion.

I. INTRODUCTION

In order to confirm a person's identity, Biometric authentication systems play a vital role. Behavioral uniqueness (signature, voice) or physiological qualities (face, iris and ear) are used for this purpose. Classification of the Biometric fusion is generally done in terms of both categories and levels [1]. Inputs or processes are defined by the categories used for fusion whereas levels used to determine how the fusion is performed. Some of the biometric fusions at different levels are given below [2]:

A. Sensor level - In this type of biometric fusion, the raw data from sensors is combined. It is referred to as image level or pixel level fusion. Sensor level fusion can be used for the advancement of multi-sample systems.

B. Feature level - The data (input) from the sensors is taken subsequent the extraction of the features at different feature sets and are formed then mutual for appropriate matching by using matching technique.

C. Matching score level - The dissimilar modules are generated, the extracted features generates the scores from dissimilar modules produce a single score by combining. The final decisions are based on Normalization and Similarity/ Dissimilarity Score.

D. Decision level - It is the highest level fusion of biometric evidences. Decision is based on the number of matches performed having threshold and it independently makes its conclusion.

E. Rank level - The output is the ranks of enrolled identities for identification. This fusion scheme is used to fuse the ranks of entity biometric systems to obtain a fused status for each individuality. It reveals less in sequence than match scores [3].

This paper is divided into following section:
Section (1) Introduction
Section (2) Objective of the proposed work
Section (3) Results and Discussion
Section (4) Conclusion

II. OBJECTIVE OF PROPOSED WORK

Feature level fusion involves the following steps for the work:

- Preprocessing of Iris, Ear and Speech at Feature level
- To propose a new algorithm for Feature extraction for iris and ear.
- Enhanced feature reduction algorithm for Iris, Ear.
- To propose a new algorithm for Feature level Fusion of Iris, Ear and Speech

The overall work of the system in Feature level fusion is shown in Figure1:

A. Preprocessing of Iris and Ear at feature level

(i) Preprocessing of Iris - Canny edge detection method is used for the enhancement of iris inner and outer boundary which is not recognized well in normal conditions. Pre-processing of the iris image, as shown in Fig.1 and Fig.2.

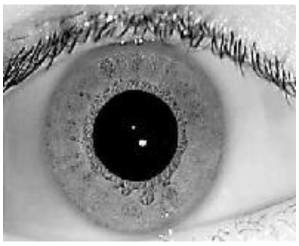


Fig 1: Eye Input image

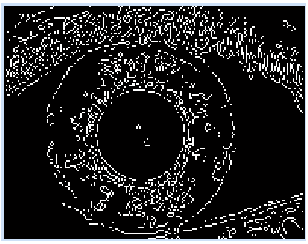


Fig 2: Canny edge detection

(ii) Segmentation - The main purpose of segmentation stage is to localize the two iris boundaries. First is the Inner boundary of iris-pupil and the other is outer boundary of iris-sclera. It is used to localize eyelids. Intensity level varies in both inner and outer parts of the pupil. Moreover pupil is darker as compared to iris. The use of Canny edge detection in pre-processing stage results in determining points of iris pupil boundary [4][5]. After determining edge points, by the use of circular Hough Transform, the center and radius of iris circle are obtained shown in Fig.3 and Fig4.

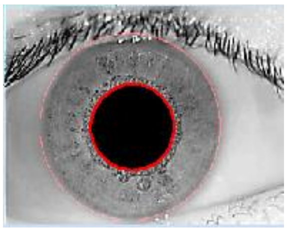


Fig 3: Segmented Iris Region

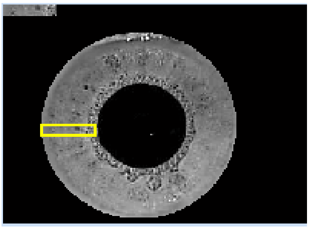


Fig 4: Iris rectangle Region

(iii) Normalization - In normalization stage, an approach based on Daugman's method is used. Iris area is transformed from polar to Cartesian coordinates.

Fig 5: Normalized Iris

It performs the mapping of each pixel in the iris into rectangular region. Therefore, iris area is obtained as a normalized strip with regard to iris boundaries and pupil center [6].

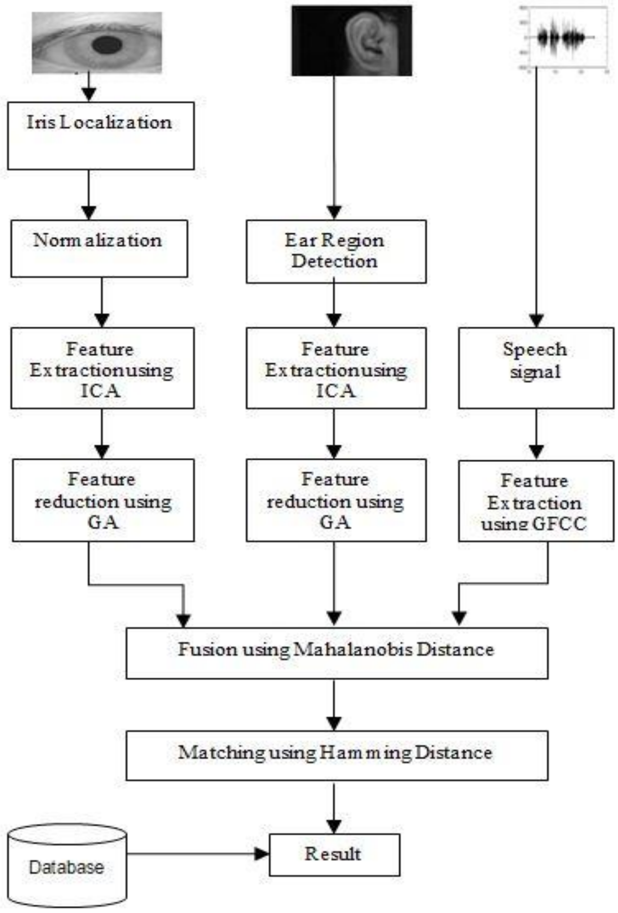


Fig 6: Flow diagram of proposed system

(iv) Preprocessing of Ear - The main purpose of image preprocessing is to remove the irrelevant information of the ear image, recovery of useful ear information, and strengthen the relevant information and Maximally simplified ear feature data, thereby improving the recognition rate [3] [6].To minimize the noise, two morphological operations are used for edge detection [7].

$$E(X, Y) = I(X, Y) - (I(X, Y) \oplus S) \Theta S \quad (1)$$

Dilation operation '$\oplus$' is used for the expands the image where erosion 'Θ' removes the redundant details from the image .Edge of image is find by using equation 1. Where E(X, Y) refers to output image, S is the structuring element for obtaining the shape information and I(X, Y) is the original image. The edge is used to form a mask. The resulting mask is multiplied with the binary form of original ear image to completely remove the areas surrounding the ear [8] [9].

Fig 7: Input Ear Image

Fig 8: Dilated Ear region

Fig 9: Detected Ear region

B. Iris and Ear Feature Extraction using ICA

(i) Preparing data - (a) Number of feature vector N where $X=(x_1,\ldots\ldots\ldots\ldots,x_n)T$ and one output class c, compose new dimension input dataset (X^T,C)^T.
(b) Normalize feature set.

(ii) Performing ICA on iris and ear features - ICA performs on new dataset and store the result matrix w of dimension

a) Shrinking small weights

1. For N independent row vector wi of w, calculate mean value ai.
2. For all w_{ij} in W, if $|w_{ij}|<b.a.a_j$, then shrink $|w_{ij}|$ to zero's is a small positive number.

b) Extracting feature values.

1. W_i, project it onto the original input feature space, delete weight w_{ic} corresponding to output class, result in new dimensional row vector. W_i'.

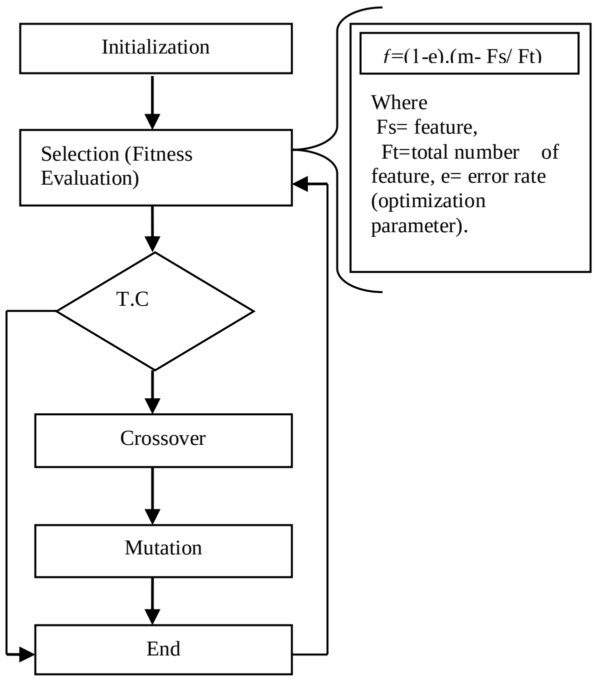
2. Multiplying new weight matrix W' of dimension(N+1)*N To the original input data X make a new vector. f_i is new feature candidates[10].

C. Iris and Ear Feature Reduction using enhanced Genetic Algorithm (GA)

The feature obtained by ICA is reduced by GA. Fitness function of GA is changed to obtain better results. In fitness function, the values calculated are based on the ranking of genetic strings. The genetic string with the global optimum could be found as the terminal clustering result by maximizing & minimizing the values. GAs is based on "natural selection and the principle of "survival of the fittest". GAs is superior as it has the ability to avoid local minima.

$$f=(1-e).(m- Fs/ Ft) \quad (2)$$

In equation (2) Where Fs= feature, Ft=total number of feature, e= error rate (optimization parameter) are used to calculate the reduced features of iris and ear [11] [12].



Flow diagram of GA

D. Speech Feature Extraction using Gammatone frequency cepstral coefficients (GFCC)

GFCC have been proposed for the extraction of feature for mandarin and English speech. Specifically, GFCC is calculated by applying a cepstral analysis on the output of Gammatone filter bank, which was initially designed according to the frequency respose of human cochlear filtering. A bank of 64 filters have been used in this experiment whose center frequency ranging from 50Hz to 8000Hz [13]. There center frequencies of Gammatone filter band are correspondingly shared along the Equivalent Rectangular Bandwidth (ERB) scale and the filters with higher center frequencies react to wider frequency ranges. GFCC, based on equivalent rectangular bandwidth (ERB) scale, has finer resolution at low frequencies than MFCC (Mel scale). Following steps show how speech signal are extracted using Gamma tone frequency cepstral coefficients (GFCC) [9][15].

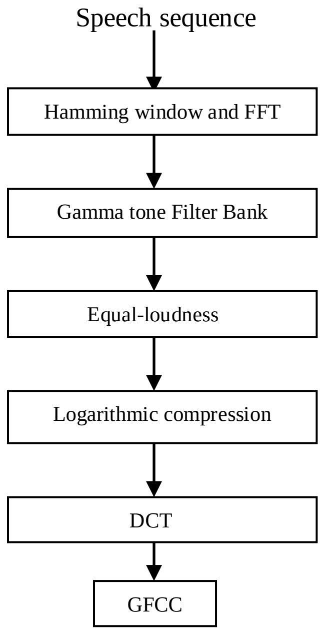


Fig 10: Flow diagram of GFCC

When some speaker information is embodied through different energy levels, the logarithmic operation better might be the case. In a noisy mixture, there are target dominant T-F units as well as segments indicative of this energy information [16].

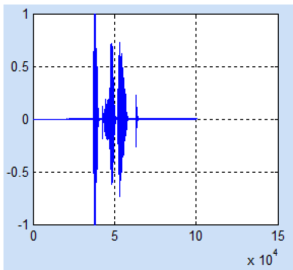


Fig 11: Original wave file

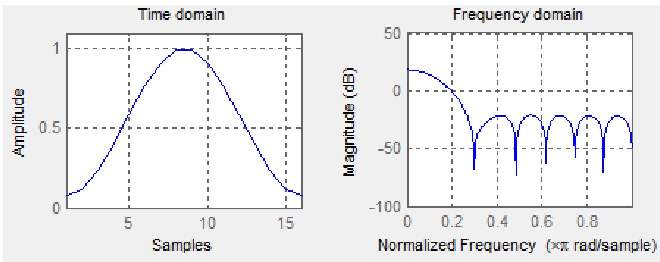


Fig 12: Hamming window, Time and Frequency domain

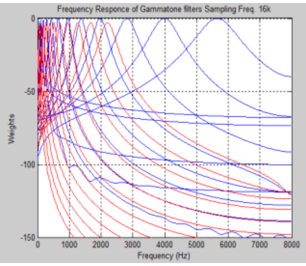


Fig 13: Gammaton filter bank

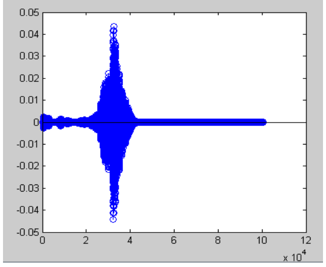


Fig 14: DCT

E. Feature Level Fusion

Feature level fusion gives the better results as compared to score level fusion .The response time to fuse the three biometrics modalities is less which gives better results. The feature vector of Iris, Ear and Speech are normalized so that they can be easily fused.

To fuse three modalities, a new method called Mahalanobis Distance is used. It is used to measure the distance between two points in the space defined by two or more correlated variables. Mahalanobis distance (Md) between a sample x, y and S is the within-group covariance matrix is calculated using following equation (1).

$$(x, y)^2 = (x - y)^t S^{-1}(x - y) \quad (3)$$

The axes in the plots would be non-orthogonal in case the variables are correlated. In this case, the simple Euclidean distance is not a good measure, while the Mahalanobis distance will be good for the correlations. After applying fusion, the mean of Mahalanobis distance is taken. Mean vector contains the average of the observations for each variable [17].

F. Matching with database

In database, the fused values of Iris, Ear and Speech are saved after fusion. Testing is done with Hamming distance [18]. In testing module feature of three modalities are extracted and then fused and matched with saved fused database.

III. RESULTS AND DISCUSSIONS

The proposed system deals with 50 person samples.Samples of the "Iris and Ear" collected from IIT Delhi database. The selected samples contain two iris images, two ear images and speech signal are manually collected.

The result obtained by Iris, Ear and Speech are explained in this section. In Iris, steps involved are: canny edge detector method used for detecting the iris inner and outer boundaries .Edge map input image is taken from Circular Hough transform and circles are drawn from the image obtained. The centre and the radius of these circles give the iris and pupil boundary.

Daugman's method is used to normalization the segmented iris region. Each point (X,Y) within iris region is mapped into polar coordinates (r, ϴ) where r is the radius and ϴ is the angle[0,2*3.14]. Normalized iris region is shown in Fig. 6.

Dilation and erosion (Morphological operations) are performed on ear image to detect the edges of ear image. Dilation residue edge detector determines the Feature extraction is done on iris and ear region using ICA which contracts small weights. GA is applied on feature extracted by ICA .GA reduces the feature vector to give better results and for fast processing. Then the feature extraction is carried out for the speech signal using GFCC .Feature band used by GFCC is 64. Nonlinear rectification is step prior to the DCT. Feature extracted from ear iris and speech is fused together. Verification is last step in which the feature vector is generated by testing image. After this it is compared with training image feature vectors .The best match gives minimum distance. It is done to improve the accuracy of biometric system and also to improve the reliability as three modalities are used. The person with fused similarity score is authenticated. Accuracy is calculated for new proposed techniques. Accuracy is calculated using false rejection rate and false acceptance rate. As shown in the

Calculate FAR – Total Number of Samples in the database=100
Number of Sample that falsely accepted = 2

$$\text{FAR} = \frac{\text{Number of Samples that falsely accepted}}{\text{Total Number of Samples}}$$

$$\text{So, FAR} = \frac{2}{100} = .02$$

Calculate FRR - Total Number of Samples in the database=100
Number of Sample that falsely accepted = 2

$$\text{FRR} = \frac{\text{Number of Samples that Falsely Rejected}}{\text{Total Number of Samples}}$$

$$\text{So, FRR} = \frac{2}{100} = .02$$

Calculate Accuracy – Accuracy = 100− (FAR+FRR) %
Accuracy = 100− (0.02+0.02) %
Accuracy = 96 %

Feature extraction level of fusion of three modalities provides more accurate results than the two modalities. The accuracy of the system is approximately 96% which is better than the other systems.

IV. CONCLUSIONS & FUTURE SCOPE

In this research, a multimodal biometric system for feature level fusion of iris, ear and speech verification system increases the accuracy of the authentication by using new technique. This system decrease the FAR and FRR respectively. The accuracy of the system is 96%.

Future works could go in the direction by using more robust modeling techniques against forgeries and hybrid fusion level can be used. Also, the system should be tested on a larger database to validate the robustness of the model by using other feature extraction methods.

V. REFERENCES

[1]. M. Ramya, A. Muthukumar and S.Kannan, "Multibiometric based authentication using feature level fusion," IEEE-International Conference on Advances in Engineering, Science and Management (ICAESM -2012) March 30, 31, 2012.

[2]. A. K. Jain, A. A. Ross, "An introduction to biometric recognition," Circuits and Systems for Video Technology, IEEE Transactions on, vol. 14, pp. 4 - 20, jan. 2004.

[3]. M.F. Nadheen and S.Poornima, "Feature Level Fusion in Multimodal Biometric Authentication System", International Journal of Computer Applications (0975 – 8887), vol 69, no.18, pp. 36-40, May 2013.

[4]. M. Mahlouji and A. Noruzi , "Human Iris Segmentation for Iris Recognition in Unconstrained Environments ," IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 1, No 3, pp.149-155,Jan. 2012.

[5]. N. Singh, D. Gandhi, "Iris recognition system using canny edge detection and a circular Hough transform," International Journal of Advances in Engineering & Technology, 2011.

[6]. M.F. Nadheen and S.Poornima, "Fusion in Multimodal Biometric using Iris and Ear Proceedings of 2013", IEEE Conference on Information and Communication Technologies, pp.83-89, 2013.

[7]. W. Zhong, "An Improved Normalization Method for Ear Feature Extraction," International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.6, No.5, pp.49-56, 2013.

[8]. Y. Hsiao ,C. Chuang , J. Jiang and C. Chien, "A Contour based Image Segmentation Algorithm using Morphological Edge Detection," IEEE International Conference on Systems, Man and Cybernetics, vol. 3 ,pp. 2962 – 2967, October 10-12, 2005.

[9]. Y.Ping and K.W.Bowyer., "Empirical Evaluation of Advanced Ear Biometrics," Proc. Empirical Evaluation Methods in Computer Vision, San Diego, (2005) 56-59.

[10]. J. Yang and D. Zhang , "Constructing PCA Baseline Algorithms to Reevaluate ICA-Based Face-Recognition Performance", IEEE Transactions on systems, man, and cybernetics—Part B: cybernetics, vol. 37, no. 4,pp. 1015-1021, Aug. 2007.

[11]. M. Satone and G. Kharate, "Feature Selection Using Genetic Algorithm for Face Recognition Based on PCA, Wavelet and SVM", International Journal on Electrical Engineering and Informatics - Vol. 6, no. 1, pp. 39-52, March 2014.

[12]. H. HsingRoad and H.Fong, "Feature selection for classification by using a ga-based neural network approach", Journal of the Chinese Institute of Industrial Engineers, Vol. 23, No. 1, pp. 55-64 ,2006.

[13]. A. Biswas, P.K. .sahu, A. bhowmick and M. Chandra, "Hindi vowel classification using gfcc and formant analysis in sensor mismatch condition", WSEAS TRANSACTIONS on SYSTEMS, Vol.13,pp. 1130-143, 2014.

[14]. K. Ajmera, S. Holambe, "Speaker Recognition Using Auditory Features and Polynomial Classifier," International Journal of Computer Applications (0975 – 8887) , , Vol. 1 – No. 14 ,2013.

[15]. M. Babu, "Whether mfcc or gfcc is better for recognizing emotion from speech ", International journal of research in computer applications and robotics, vol.2 issue.6, pg.: 14-17,june 2014

[16]. E. B. Tazi, A. Benabbou and M.Harti, "robust features for noisy text-independent speaker identification using gfcc algorithm combined to vad and cmntechniques", Journal of Theoretical and Applied Information Technology**",** vol. 36 no.2, pp. 206-216, Feb. 2012.

[17]. U. Gawande, M. Zaveri,and A. Kapur, "A Novel Algorithm for Feature Level Fusion Using SVM Classifier for Multibiometrics-Based Person Identification", Hindawi Publishing Corporation Applied Computational Intelligence and Soft Computing , Article ID 515918, pp. 1-11,2013.

[18]. R. Mishra and V. Pathak, "Human Recognition using Fusion of IRIS and Ear Data", International Conference on Methods and Models in Computer Science , 2009.

[19]. L. P. Rodriguez, A. G. Crespo, "Study of Different Fusion Techniques for Multimodal Biometric Authentication," in Networking and Communications, 2008. wimob '08. IEEE International Conference on Wireless and Mobile Computing, 2008, pp. 666 -671.

[20]. N B. Boohoo, R K. Subramanian, " Robust multi-biometric Recognition Using Face and Ear Images, "International Journal of Computer Science and Information Security, Vol. 6, No. 2, , 2009.

Author Profile

Harpreet Kaur is working as an Assistant Professor in CGC Technical Campus, Jhanjeri (Mohali).She received the B.Tech degree in Information Technology from Baba Banda Singh Bahadur Engg. College, Fatehgarh Sahib during 2008-2012 and M.Tech degree in Computer Science Engineering from Sri Guru Granth Sahib World University, Fatehgarh Sahib during 2012-2014 respectively.

Ranjit Singh is working as an Associate Professor in CGC Technical Campus, Jhanjeri (Mohali), Punjab. He received his M.Sc. in Mathematics from Panjab University, Chandigarh, India in 2000 and M.Phil. in Mathematics in 2004. He is also pursuing his Ph.D. in Mathematics from Punjab Technical University, Jalandhar, Punjab, India. He is CSIR-UGC NET qualified. He has over 12 years of teaching experience of teaching Engineering mathematics, quantitative techniques and operations research. He has many publications in national and international Journals. He is reviewer for many international Journals. His field of interest is Operation Research (OR) and its applications in financial analysis.