ISSN:2249-5789
M Pujitha Raj et al , International Journal of Computer Science & Communication Networks,Vol 5(2),67-71

# Multi-modal Biometric system using ear and face(2D+3D) Modalities

M.Pujitha Raj
Computer Science and engineering
Amrita University
Coimbatore, India
poojithampalli@gmail.com

Manjusha.R
Computer Science and engineering
Amrita University
Coimbatore, India
R_manjusha@cb.amrita.edu

B.Achyut Sarma
Computer Science and engineering
Amrita University
Coimbatore, India
achyutsarmaboggaram@gmail.com

S.Vaishnavi
Computer Science and engineering
Amrita University
Coimbatore, India
vaishnavi.sunkara93@gmail.com

***Abstract*— We have made a research on the field of biometrics and found that multimodal biometric systems plays a major role in authentication and security for different types of areas such as science, medicine, engineering, robotics, etc. Face recognition using biometrics is the frontier of the security systems. The ear, which is a biometric identifier, that has some desirable properties such as universality, uniqueness, permanence, can also be used for security purposes. We propose a multi-modal biometric system combining face (2D+3D) and ear features at various levels which uses Microsoft Kinect. Commencing with the survey of existing algorithms, applied to face (2D+3D) and ear data, we focused on fast discrete curvelet transformation techniques for face (2D+3D) recognition and active contour algorithm techniques for ear recognition. Finding optimal fusion level and avoiding redundancy in the extracted features are some challenges in designing our system. Our system is insensitive to lighting conditions, pose variations, aging and can completely replace the current recognition systems economically and provide a better security. A total of 250 subjects participated in data acquisition sessions through kinect. The results are obtained separately each biometric and fused at metric level for higher accuracy. Our multimodal algorithm performed better by achieving 97 percent and 95 percent verification rates with 0.01 false acceptance rate which is greater than either 2D face or 3D face or ear recognition algorithm alone in a statistical and significant manner.**

***Keywords—Biometrics; Security System; Person Identification; Modalities; Kinect; Feature Extraction; Curvelets; Eigen Faces; Geometric Features; Feature level Fusion.***

## I. Introduction

Biometrics is defined as the unique physical characteristics or traits of human body. The details of human body which will be used as unique biometric data to serve as ID, such as: retinal, iris, fingerprint, palm print and DNA. Since these biometric systems provide superior level of person identification which shows a key point about identification that is “inalienable” means no subject can be taken away from or given away by the possessor, so no more hackings can be done to authentication systems – they have to go through the control point personally [1].

Many researchers have concluded that multimodal biometrics systems have more advantages over unimodal biometric systems. Multimodal biometric systems have become the best suited solution for any industry where high accuracy is required because the integration of two or more types of biometric verification systems helps to meet stringent performance requirements. Face (2D+3D) and ear modalities can produce absolutely unique data sets when done properly. It is not easy to find the best method for data acquisition, each of the different methods has inherent advantages and disadvantages.

Face 2D authentication is based on curvelet transform and PCA we decompose face images to get low frequency coefficients by curvelet transform. PCA with an exponential decay factor is applied on these selected coefficients to extract features vectors, which will achieve dimension reduction as well. As each biometric technology has its own merits and demerits, these systems have difficulty in recognizing a face from images captured from two drastically different views and under different illumination conditions [2].

Moreover, to improve the accuracy rate of the system to a higher degree level we have combined ear recognition with face recognition to endow “multimodal system” that guarantees accuracy through multiple layers of security. In the need of enhanced security, 3D face verification can be used in integration with face 2D and ear verification systems which would also keep a log of 3D facial data which could be used as evidence [3].

This paper proposes a multi-modal biometric system for authentication that fuses the features of face (2D+3D) with that of the ear for increasing authentication accuracy rate. The feature sets best performing for each biometric are fed to a feature fusion strategy for increasing authentication accuracy.

The rest of the paper is organized as follows. Section 2 scans the previous work in the related areas. Section 3 presents the structure of the proposed system. Section 4 analyzes the obtained results. The paper is terminated by a conclusion summarizing the obtained results and specifying the problems for future work.

## II. Previous Works

S.Elaiwat, M.Bennamoun, F.Voussaid, and A.El-Sallam[3] presented a robust single modality feature-based algorithm for 3D face recognition. The proposed algorithm by them exploits Curvelet transform not only to detect salient points on the face but also to build multi-scale local surface descriptors that can capture highly distinctive rotation invariant local features around the detected key points. There approach is show to provide robust and accurate recognition under varying illumination conditions and facial expressions. It is not easy to predict how a biometric technology could be evolved and get embedded in which applications. But it is shown that biometric-based recognition will have a profound influence on the way we conduct our daily business.

Alaa ELEYAN, and Hasan DEM˙IREL [5], introduced a new face recognition technique based on the gray-level co-occurrence matrix (GLCM). GLCM represents the distributions of the intensities and the information about relative positions of neighboring pixels of an image. They proposed two methods to extract feature vectors using GLCM for face classification. The first method extracts the well-known Haralick features from the GLCM, and the second method directly uses GLCM by converting the matrix into a vector that can be used in the classification process. The results demonstrate that the second method, which uses GLCM directly, is superior to the first method that uses the feature vector containing the statistical Haralick features in both nearest neighbor and neural networks classifiers. The proposed GLCM based face recognition system not only outperforms well-known techniques such as principal component analysis and linear discriminant analysis, but also has comparable performance with local binary patterns and Gabor wavelets. It is obvious from the results that the GLCM is a robust method for face recognition with competitive performance.

Ping Yan and Kevin W. Bowyer [6] works have shown that the ear is a promising candidate for biometric identification. The preprocessing of ear images has had manual steps and algorithms have not necessarily handled problems caused by hair and earrings. They presented a complete system for ear biometrics, including automated segmentation of the ear in a profile view image and 3D shape matching for recognition and evaluated the system with the largest experimental study to date in ear biometrics, achieving a rank-one recognition rate. Future work can be on the effect on pose on ITC matching results and to examine weather eyeglasses can cause a shape variation in the ear. Further study should in guidelines that provide best practices for the use of 3D images for biometric identification in production systems.

Ajmal S. Mian, Mohammed Bennamoun, and Robyn[7] Owens presented a fully automatic face recognition algorithm and demonstrated its performance on the FRGC v2.0 data. Their algorithm is multimodal (2D and 3D) and performs hybrid (feature based and holistic) matching in order to achieve efficiency and robustness to facial expressions. The pose of a 3D face along with its texture is automatically corrected using a novel approach based on a single automatically detected point and the Hostelling transform. A novel 3D Spherical Face Representation (SFR) is used in conjunction with the Scale-Invariant Feature Transform (SIFT) descriptor to form a rejection classifier, which quickly eliminates a large number of candidate faces at an early stage for efficient recognition in case of large galleries. The remaining faces are then verified using a novel region-based matching approach, which is robust to facial expressions. Their approach automatically segments the eyes, forehead and the nose regions, which are relatively less sensitive to expressions and matches them separately using a modified Iterative Closest Point (ICP) algorithm. The results of all the matching engines are fused at the metric level to achieve higher accuracy. They used the FRGC benchmark to compare our results to other algorithms that used the same database.

Mahoor, Denver, Cadavid and Abdel-Mottaleb[4] described a multi-modal ear and face biometric system. The system is comprised of two components: a three-dimensional (3D) ear recognition component and a two-dimensional (2D) face recognition component. For the 3D ear recognition, a series of frames is extracted from a video clip and the region of interest (i.e., ear) in each frame is independently reconstructed in 3D using Shape From Shading. The resulting 3D models are then registered using the Iterative Closest Point algorithm. They iteratively considered each model in the series as a reference model and calculated the similarity between the reference model and every model in the series using a similarity cost function. Cross validation is performed to assess the relativity of each 3D model. The model that demonstrates the greatest overall similarity is determined to be the most stable 3D model and is subsequently enrolled in the database. For the 2D face recognition, a set of facial landmarks is extracted from frontal facial images using the Active Shape Model technique. Then, the responses of the facial images to a series of Gabor filters at the locations of the facial landmarks are calculated. The Gabor features (attributes) are stored in the database as the face model for recognition. The similarity between the Gabor features of a probe facial image and the reference models are utilized to determine the best match. The match scores of the ear recognition and face recognition modalities are fused to boost the overall recognition rate of the system.

## III. Proposed System

Since there are high requirements in security issues, biometric authentication systems have been commonly utilized in many recognition applications. Multimodal systems have great demands to overcome the issue involved in single trait systems and this has become one of the most important research areas of face and ear recognition. This paper presents

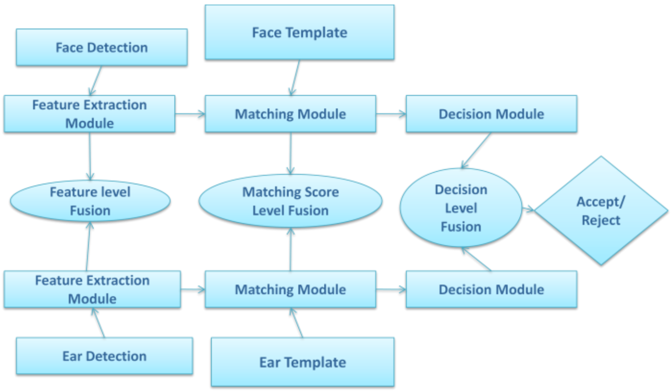



Figure.1 Architecture of proposed system

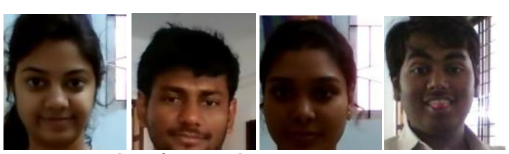
a multimodal face(2D+3D) and ear biometric verification system to improve the performance that fuses ear features and face features for better authentication accuracy. The architecture for the complete system is shown in figure 1. The proposed system proceeds as follows:

### A. *Preparation phase*

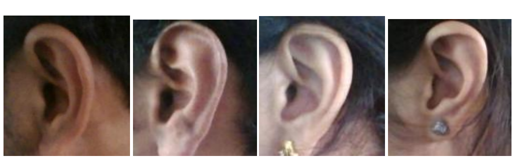
In this phase, we estimate the performance of four statistical feature extraction methods and adopt the one with the best recognition rate. The process proceeds as follows shown in figure.3:

1. *Image Aqcquisition*

The face 2D images are captured using Microsoft Kinect in color sensor mode for 250 subjects as shown in figure 2(a).
The face 3D images are captured using Microsoft Kinect in depth sensor mode for 250 subjects as shown in figure 2(b).
The ear modality images are also captured using Microsoft Kinect in general mode for 250 subjects as shown in figure 2(c).



Examples of captured 2D images using Kinect



Examples of captured ear images using Kinect

Examples of captured depth images using Kinect

Figure.2 Examples of captured images of face(2D&depth) and ear

2. *Preprocessing*

Several processing techniques are used to enhance the image quality. On the other hand, facial data images were noise-clean and contrast enhanced.

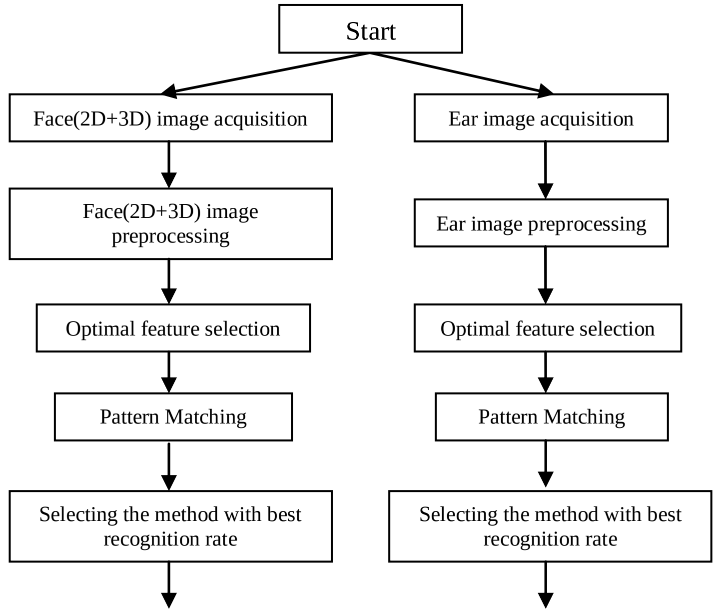



Figure.3 Flow process of preparation phase.

3. *Feature Selection*

The following approaches are applied for feature extraction; namely; statistical features (SF), movement invariants (MIs), multiscale local surface descriptors, curvelet coefficients and a robust feature extraction algorithm for ear. The paper aims at selecting the best performing approach for both biometrics. The used feature selection techniques are as follows:

- *Statistical features*

  The following set of features is used Mean,variance,smoothness,skewness, uniformity and entropy.
- *Moment invariants:*

  The most frequently used invariants moments which are invariant under cancellation changes in scale and also rotation.
- *Multiscale local surface descriptors*
- *Curvelet coefficients*
- *Robust feature extraction algorithm for ear:*

  The phases of this algorithm include skin region detection, surface curvelet estimation, surface segmentation and classification.

### B. *Training phase*

The training phase is depicted in figure.4. In this phase:

- The face and ear images are captured in the same way as preparation phase.

ISSN:2249-5789



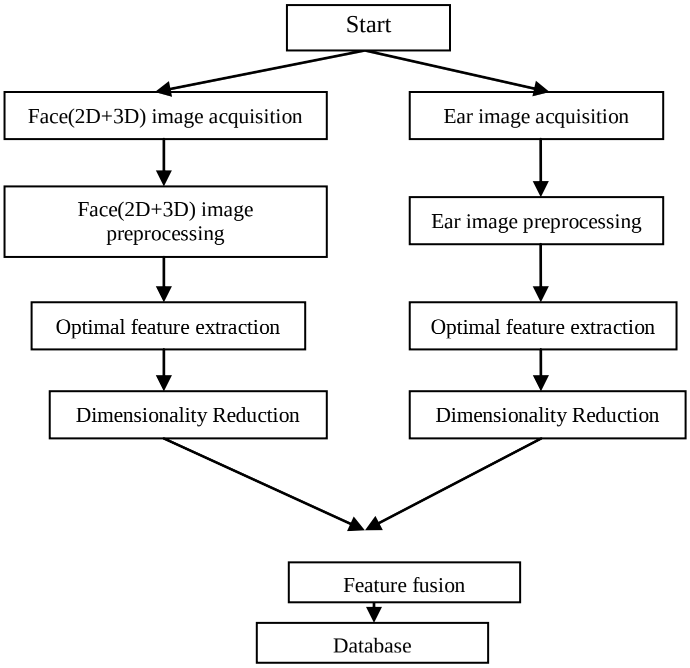



Figure.4 Flow process of training phase.

- The captured images pass through the same preprocessing stages.
- The feature vectors are extracted using appropriate technique.
- Dimensionality reduction is performed on both the feature vectors.
- The feature vectors of both biometrics are fused using techniques such as Max-score, Min-score, and Product-of-score.
- The fused feature vector is stored in a database for future comparisons.

### C. *Testing phase*

The testing phase is depicted in figure.5.

- The images of the person under test are acquired and preprocessed typically as in preparation phase.
- The feature vector of the adopted approach is calculated and reduced.
- The resulting feature vector is compared with those stored in database and person is recognized.

## IV. Results And Discussion

Fig.6 illustrates the results of the verification system. The results are presented as Receiver Operating Characteristic curves (ROC) with the fusion of all the modalities. As the ROC curve demonstrates, the ear and face (2D+3D) modalities have a verification rate of 95%, 75% and 97% at .01 False Acceptance Rate(FAR), respectively.

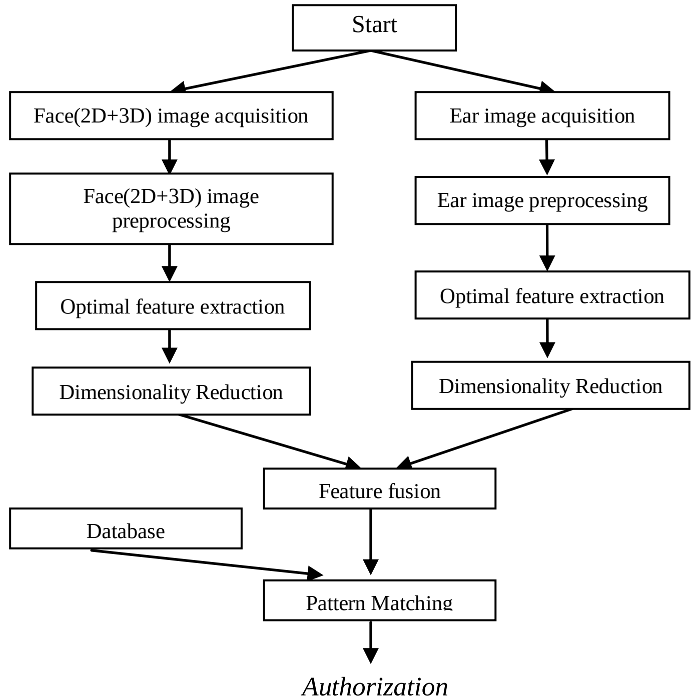



Figure.5 Flow process of testing phase.

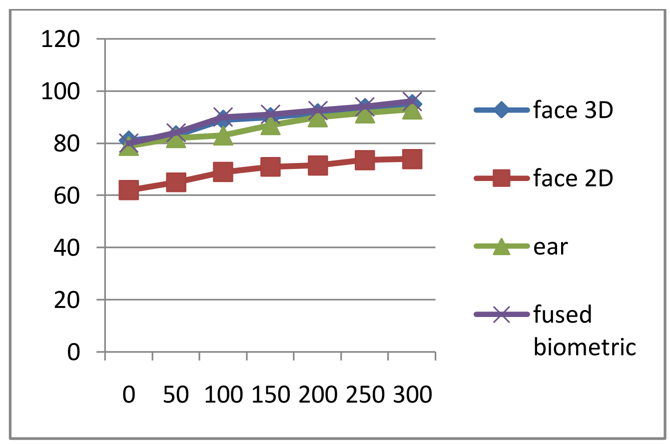



Figure.6 Cumulative match characteristic curves of face (2D+3D), ear and fusion of three modalities representing Rank in x-axis and Identification rate(%) in y-axis.

| Fusion Technique | Identification(%) | Correct verification rate @.01FAR |
|---|---|---|
| Max-Score | 95 | 95.1 |
| Min-Score | 81.2 | 79.3 |
| Product-of-Score | 97.9 | 97.9 |
| Weighted Sum | 98 | 98 |

Table.1 Different te'chniques used to fuse the normalized match

ISSN:2249-5789


We have also investigated the use of other techniques for data fusion (i.e., Max-Score, Min-Score, and Product-of-Score) and compared their results with the weighted sum technique. Table.1 compares the rank-one identification rate, and the correct verification rate at .01% FAR of the other techniques for fusion. As this Table illustrates, the weighted sum technique outperforms the other techniques for data fusion.

| Noise intensity | Recognition rate using feature vectors of | | |
|---|---|---|---|
| | Face 2D | Face 3D | Ear |
| 0% | 100% | 100% | 100% |
| average | 75% | 97% | 95% |

*Table.2 Recognition rate of images corrupted by average noise using face 2D,3D, ear and fused feature vectors.*

For evaluating the performance of the proposed system, we calculated the recognition rate images corrupted by noise with average intensity . The obtained results are shown in Table 2.

## V. Conclusions and Future work

This paper presents how ear and face(2D+3D) modalities are fused at different fusion levels and found maximum accuracy at the match score level. The images under consideration are corrupted by different types of noise with different noise intensities, and the recognition rate is evaluated for the average case. It was found that Moment Invariant (MIs) feature vector guarantees the best recognition rate in all cases.

For face recognition curvelet transform is used to extract the feature vector coefficients and for ear active contour is used to extract the features. The extracted features were then used to calculate the similarity between images. This resulted in match score for each modality. The match scores obtained from the two modalities (ear and face) were fused at the match score level using the weighted sum technique.

Future work will include the employment of 3D face recognition in combination with 3D ear recognition with 100% accuracy. The use of the 3D modality for both biometric markers will lead to an increase in the robustness to both illumination and pose variations.

## References


[1] K.Sasidhar, Vijaya L Kakulapati, Kolikipogu Ramakrishna & K.KailasaRao, "Multimodal biometric systems-Study to improve accuracy and performance," Internatioal journal of Computer Science & Engineering Survey (IJCSES) Vol.1, No.2, November 2010.

[2] Ma Hui, Hu Fengsong, "Face recognition using Curvelet Transform and (2D)2 PCA," I.J. Education and Management Engineering 2012,6,23-28, July 2012 in MECS.

[3] S.Elaiwat, M.Bennamoun, F.Boussaid, and A.El-Sallam, "3-D Face Recognition using curvelet local features," IEEE Signal processing letters, VOL. 21, NO. 2, FEBRUARY 2014.

[4] Mahoor, Denver, Cadavid and Abdel-Mottaleb, "Multimodal ear and face modelling and recognition," IEEE Image Processing,2009.

[5] Alaa ELEYAN, Hasan DEMIREL, "Co-occurance matrix and its statistical features as a new approach for face recognition," Turk J Elec Eng & Comp Sci, Vol.19, No.1, 2011.

[6] Ping Yan and Kevin W. Bowyer, Fellow, "Biometric Recognition Using 3D Ear shape," IEEE transactions on PATTERN ANALYSIS AND MACHINE INTELLIGENCE, vol. 29, No. 8, August 2007.

[7] Ajmal S. Mian, Mohammed Bennamoun, and Robyn Owens, "An efficient Multimodal 2D-3D Hybrid Approach to automatic Face Recognition," IEEE transactions on pattern analysis and machine intelligence,Vol.29,No.11,November2007.

[8] S.F.Bahgat, S.Ghoniemy, M.Alotaibi, "Proposed multimodal palm veins-face biometric authentication," International journal of advanced computer science and applications, Vol.4, No.6, 2013.

[9] Priyanka Dhoke, M.P.Parsai, Amatlab based face recognition using PCA with back propagation neural network," International Journal of innovative research in computer and communication engineering, Vol.2, Issue 8, August 2014.

[10] Lin Zhang, Zhixuan Ding, Hongyu Li, Ying Shen,"3D Ear identification based on sparse representation",PLoS ONE, march 2014.

[11] Jiamin LIU, Fulin LUO, Hong HAUNG, Yizhe LIU, "Ear detection based on improved camshift and adaboost algorithm", Journal of computational information systems 10: 13, 2014.

[12] W.Zhao, R. Chellappa, P.J Phillips, A. Rosenfeld, "Face Recognition: A Literature Survey," ACM Computing Surveys, vol.35, no.4, pp.399-458, 2003.

[13] K.W.Bowyer, K.Chang, P.Flynn, "A survey of approches and challenges in 3D and multimodal 3D+2D face recognition," Computer vision and image understanding, vol.101, Issue 1, pp.1-15,

[14] P.J.Phillips, P.J. Flynn, T. Scruggs, K.W. Bowyer, J. Chang, K.Hoffman, J.Marques, J.Min, W. Worek, "Overview of the face recognition Grand Challenge," Computer vision and pattern recognition, IEEE Computer Society Conference, Vol.1, pp.947-954, 2005.