

# Multimodal Database: Biometric Authentication for Unconstrained Samples

S. Poornima

**Abstract** Biometrics provides a reliable authentication of a human in a wide variety of applications such as security systems, surveillance and human-computer interaction. Biometric system was started with utilization of a single biometric feature referring as a unimodal biometric system, which is unable to fulfill the security needs extensively in a highly sensitive environment and hence multibiometrics has emerged gaining more importance in the research area. Though there is a shortage of publicly available multimodal databases acquired in real unconstrained environment, a multimodal biometric system can succeed with the assistance of suitable multiple sensors providing higher accuracy rate than that of unimodal biometrics, of course subject to cost, time and subject's acceptance This paper presents a new multimodal dataset which is developed using simple acquisition setup and devices to capture features belonging to the same person in uncontrolled scenarios. The dataset is composed of color images collected from 100 subjects (50 male and 50 female) under the age group 18–22. Totally 6 samples per trait were collected at different time internals between 2011 and 2014 with various occlusions. The dataset is also tested and analyzed by our developed biometric recognition system.

**Keywords** Biometrics • Multimodal • Segmentation • Noise removal • Feature extraction • Fusion

## 1 Introduction

Biometric serves society in various real time security applications, providing high recognition rates and low error rates. Unibiometric systems suffer from limitations such as non-universality, noisy sensor data, large intra user variations and susceptibility to spoof attacks [1], whereas multibiometrics [2, 3] addresses the above issues and provides improved performance with very low error rates. However,

---

S. Poornima (✉)  
SSN College of Engineering, Chennai, India  
e-mail: poornimas@ssn.edu.in

there is immense opportunity for betterment and it must be acceptable to general public for being user friendly, convenient and affordable for having satisfactory level of reliability and security. But researchers run the risk of high cost and high storage space when using multibiometric system. Consequently, there is a shortage of publicly available multimodal databases acquired in real unconstrained environment and most of the systems are experimented on virtually built multimodal databases from publicly available unimodal datasets either into signals or scores. As well as, when multiple traits are involved, acquisition of input samples requires different sensors like scanners, camera or printer and further, the samples of multiple features traits of the same person is to be acquired. Under unique circumstances, the user must not feel discomfort and annoyed during acquisition of inputs using different sensors and certain constraints like position, distance, and contact nature. Taking these into consideration along with two essential characteristics of a biometric system, namely collectability and acceptability, our objective is to promote higher collectability and choice of feature which helps continue its significant growth in security related applications. Hence the above things motivate us to create a new database using a cost effective, suitable acquisition device such as any digital or mobile camera to capture contactless features like iris, face, ear and a web camera to acquire Dorsa palm vein of a person respectively. These features are chosen to carry out research work in the area of multibiometrics, which is suitable and adaptable for any real time biometric applications.

This paper explores the impact of various biometric features on our developed biometric systems. In Sect. 2, the proposed system that handles iris, face, ear and palm dorsal vein characteristics is presented. Section 3 describes the experimental analysis and results obtained by matching individual biometric traits and concluded in Sect. 4.

## 2 Description and Imaging Principle of Multimodal Dataset

The type of biometric features considered in this proposed system requires different and suitable acquisition devices. Conventionally, high quality scanner for iris, digital camera for face and ear and an infrared camera for vein are used to capture these features. Our intention is to introduce a reliable, versatile and economical acquisition setup to collect the above discussed biometric features and also suitable for small scale biometric applications to provide security access. The noisy samples are collected from 100 subjects (50 Males and 50 Females) from the age group of 18-30, totally 6 samples per trait, in different intervals of 2011–2014, to account for time variance. A digital/mobile camera of not less than 5 Megapixels is used to acquire iris, face and ear in an unconstrained environment. The distance of the camera is variable (20–50 cm) depending on the type of image sample to be acquired. The datasets contain various samples of normal condition and with



**Fig. 1** Palm vein acquisition setup; *Top view; Array of LEDs (Left to right)*

occlusions such as eyelids drooped, half closed eye, reflections, shadows, wearing contact lens, scarf, sunglasses, hazy samples, ornaments, covered with hair, illuminations, and blurred samples, etc.

An INTEX WEBCAM IT-LITE-VU is used as a Near Infrared (NIR) camera to capture Dorsa palm vein patterns of human and the setup is shown in Fig. 1 [4]. INTEX WEBCAM IT-LITE-VU has 1/7" CMOS sensor with a frame rate of 30fps and its focus distance ranges from 4 cm to infinity. The lens view angle is around  $54^\circ$ . The camera produces an image around 15 megapixels. This camera is designed for taking images in the visible spectra by blocking out the infrared light using an IR filter. A camera is converted into an IR camera by removing the IR filter and placing a filter for visible light. The best filter for visible light is to use a new negative photographic film, which blocks out visible light and allows Infrared light to pass through the camera. To view the vein patterns under a near infrared camera we need an infrared source emitting infrared rays in near-infrared region. This illuminates the underlying vein patterns and can be viewed under the near-infrared camera.

The vein patterns are viewed using 30 infrared LEDs connected serially in a breadboard powered by 18 V battery source with precision of 780 nm light source. A black background is chosen to improve the perspective of the acquired hand images. The camera is mounted horizontally parallel to the base on which the hand is placed at a height of 34 cm. The array of infrared LEDs emits infrared rays in all directions. In order to regulate the amount of light falling on the hand, a breadboard is placed in an angle of  $60^\circ$  to the platform on which the camera is mounted. A hand is placed on a slope at an angle of  $50^\circ$  to the base to provide acute focus on our region of interest which includes the knuckle tips and the surface of the dorsal palm. All acquired images are processed to fix to a black background for visual clarity of the feature [4]. Image samples of the database are shown in Fig. 2.



Fig. 2 Samples of biometric traits from multimodal database

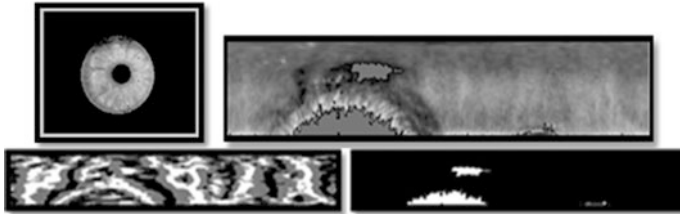
### 3 Overview of Unimodal Biometric Systems

The proposed biometric system is developed for different physiological biometric features, namely, iris, face, ear and Dorsa palm vein. A biometric system consists of four main modules such as acquisition, preprocessing, feature extraction and matching. In the following subsections, we focus individually on the processing methods of each distinct feature, namely iris, face, ear and Dorsa palm vein of the same individual for the samples collected in the above multimodal database.

#### 3.1 Processing of Iris

Images captured in an unconstrained environment contain various noises. Before starting the process of the collected features, the sample images are preprocessed using median filter and Gaussian filter to reduce or remove noise, according to the type of noise exists in the input sample. The main modules in the processing of iris feature are iris localization and feature extraction [5–8].

Since the iris region is surrounded by outer sclera and inner pupil boundaries, iris is segmented by detecting the edges of iris as well as pupil using Roberts and Canny edge detectors respectively [5]. Circular Hough transform helps to detect the centre and radii of both iris and pupil boundaries using the property of a circle. Followed by, unrelated parts like eyelid and eyelashes which acts as noise are also to be removed [5]. Iris normalization compensates the stretching of texture according to the changes in pupil size and maintains the same texture information, regardless of pupil dilation [6]. Normalization approach produces a 2D array using the pixel coordinates within the iris region which is performed using Daugman's rubber sheet model [6, 8] by mapping of each pixel in the iris region into rectangular region like unwrapping the image considering the size inconsistency and dilation of pupil and



**Fig. 3** Noise removed iris; Normalized representation; Feature template; Noise mask (*Left to right*)

generates a normalized representation of iris pattern. But the rotational inconsistencies are handled using shifting operation as in Daugman’s system.

The discriminant information from the normalized iris are extracted using Log-Gabor filter, where zero DC components can be obtained for any bandwidth on a logarithmic scale [9]. Log-Gabor filter in terms of frequency response is given as,

$$G(f) = \exp\left(\frac{-\log\left(\frac{f}{f_0}\right)^2}{2\log\left(\frac{\sigma}{f_0}\right)^2}\right). \tag{1}$$

where,  $f_0$  is the centre frequency, and  $\sigma$  represents the bandwidth of the filter. Convolve the filter with the normalized iris image for the given the wavelength and bandwidth and produce the iris template and a noise mask relatively [5]. Sample outputs from iris processing is shown in Fig. 3.

### 3.2 Processing of Face

The input image is initially cropped to have focus. This image is preprocessed using Gaussian filter to remove any noise and ROI of input, i.e., face region is extracted



**Fig. 4** Preprocessed face samples

using a robust face detection method, Viola-Jones algorithm [10] and which helps to detect feature points in the image. The features from the face region detection are extracted using a Gabor function similar to iris. Gabor function is one of the prominent functions used for extraction of the feature points in an image. Using the Gabor wavelets generated by this function we obtain the base features for comparing any input images (Fig. 4).

### 3.3 Processing of Ear

An ear is an external biometric feature which is easier to acquire and process in two steps namely ear detection and feature extraction. Noises in the image would degrade the edge detection result. Hence the noise effect was removed by convolving the image with Gaussian operator. Morphological operations [11] are used to analyze the shape of the image by choosing an appropriate structuring element. The primary morphological functions, dilation and erosion process are performed by laying the structuring element B on the image F and sliding it across the image in a manner similar to convolution [11]. The edges of the image are determined by the dilation residue edge detector. Dilation allows objects to expand, thus potentially filling in small holes and thus connecting the disjoint objects. Erosion shrinks objects by fetching away (eroding) their boundaries. This morphological operator helps to find the clean edges of the input ear using a sample as a structuring element. It is observed that inner ear region possesses a higher amount of information useful for recognition (Fig. 5). The essential features are extracted from the detected ear region using a shape descriptors.

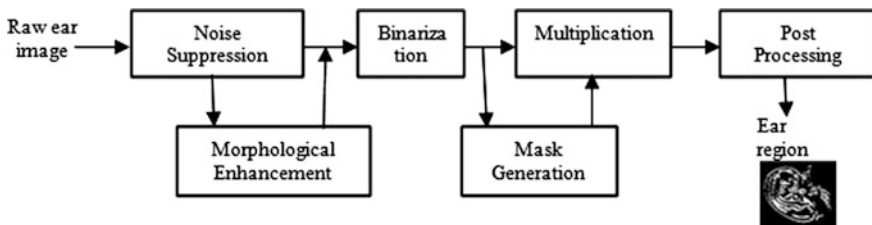


Fig. 5 Framework for ear detection

### 3.4 Processing of Palm Vein Pattern

The images acquired from the web camera are of dimensions  $640 \times 480$ . The image acquired is reduced to a dimension of  $320 \times 240$  so as to improve the visibility of the vein patterns and is presented for further processing. The images are then passed through a median filter initially for removing noise, followed by a low pass and a Gaussian filter for image smoothing. Thereafter, improved the contrast of the image by Contrast-Limited Adaptive Histogram Equalization and binarized using Otsu threshold. On this binarized image, canny method is applied for detecting edges and the image is cropped to locate the region of interest. These images are again filtered and applied with adaptive threshold. Further, the obtained output sample is then thinned to provide the final output region of interest (Fig. 6) and features are extracted using Hierarchical Multiscale Local Binary Pattern (HMLBP) which helps to extract the non uniform patterns present in the ROI and enhances the depth of information through feature extraction.

$$H(k) = \sum_{i=1}^N \sum_{j=1}^M f(\text{LBP}_{P,R}(i,j), k), \quad k \in [0, K]. \tag{2}$$

Where,

$$f(x, y) = \begin{cases} 1, & x = y \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

$$\text{LBP}_{P,R} = \sum_{P=0}^{p-1} s(g_P - g_C) 2^P. \tag{4}$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{5}$$

where  $K$  is the maximal LBP pattern value [12]. In Multi-scale LBP [12], histogram is built from bigger to smaller radius and grouping the patterns as, ‘uniform’ and

**Fig. 6** Sample extraction of region of interest



'non-uniform' through mapping each point. Hence, a sub histogram is constructed from the uniform patterns and LBP patterns from non-uniform pattern are extracted starting smaller radius. Finally, these histograms are concatenated into one multi-scale histogram [13].

### 3.5 Distance Matching

Hamming Distance (HD) is used for the measurement of similarity distance between the test and training templates of iris, ear and vein. A decision can be made to determine the relation between the templates whether they belong to the same feature or different one. Sometimes there will be an overlap between inter-class and intra-class distributions which would result in higher false acceptances and rejections. Hence, two thresholds (T1, T2) have been set. One for intra and other for inter class comparisons respectively. An intra class comparison is determined as a match if the HD measure is  $\leq T1$ , and if  $< T2$ , but  $> T1$ , it determines that probability of having images belong to the same feature [5]. The modified HD is defined as,

$$HD = \frac{1}{N - \sum_{k=1}^N X_{nk} \text{ (OR)} Y_{nk}} \sum_{j=1}^N X_j \text{ (XOR)} Y_j \text{ (AND)} X_{n'j} \text{ (AND)} Y_{n'j}. \quad (6)$$

where  $X_j$  and  $Y_j$  are the two bit-wise templates,  $X_{n_j}$  and  $Y_{n_j}$  are the corresponding noise masks for  $X_j$  and  $Y_j$ , and  $N$  is the number of bits represented by each template [5, 7]. The similarity distances are measured for all three feature types by comparing their respective trained templates. The lowest score among each type gives the better match of it.

## 4 Experimental Setup and Results

Two samples from each trait have been trained and the rest of the samples are utilized for testing. There are certain parameters to be adjusted that influence the performance of the system. The parameters used in the normalization phase of iris feature and used in feature encoding phase of all traits had to be adjusted to give maximum performance. These parameters are selected using the decidability factor [7], which is a

**Table 1** Thresholds for intra and inter class comparisons

Feature type	T1	T2
Iris	0.349	0.443
Ear	0.17	0.25
Palm (Dorsa) vein	0.277	0.391
Face	0.315	0.392



**Table 2** Performance measures of multimodal dataset

Feature type	FAR	Accuracy
Iris	0.1	89
Ear	0.1	88
Palm vein	0.1	87.3
Face	0.1	82.3

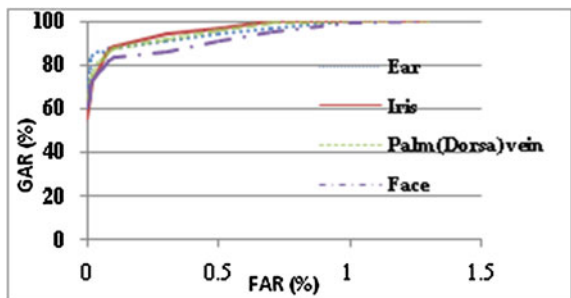
function of mean and standard deviation of intra and inter class comparisons. The higher the decidability, the greater is the separation of intra-class and inter-class distributions [5]. The computed distance measures are analyzed based on thresholds T1 and T2 (Table 1) to find out the match decision. The overall performances obtained for multimodal dataset are given in Table 2. Individual measures of each trait at various thresholds are shown in Table 3 and Fig. 7. The proposed system is also experimented with few existing public dataset whose performances are given in Table 4.

The accuracy rates obtained for samples of multimodal dataset is moreover equivalent to the performance obtained for public dataset samples. When multiple features are considered for recognition, the features must belong to the same person for an exhausted analysis. Hence the fusion at score level is implemented by

**Table 3** Performance measures of biometrics traits at various thresholds

Iris			Face			Ear			Palm (Dorsa) vein		
T1	T2	GAR	T1	T2	GAR	T1	T2	GAR	T1	T2	GAR
0.10	0.20	40.7	0.10	0.19	65.1	0.10	0.18	78.46	0.08	0.18	68.5
0.16	0.26	73	0.12	0.20	69.9	0.12	0.20	82.3	0.11	0.22	72.1
0.20	0.31	75.3	0.14	0.22	72.8	0.13	0.21	84.3	0.14	0.25	75.2
0.25	0.34	80	0.16	0.29	76.1	0.14	0.22	85.9	0.18	0.30	80.7
0.29	0.39	84.6	0.21	0.29	78.7	0.15	0.23	86.3	0.22	0.33	82.6
0.30	0.40	86.3	0.27	0.35	80.5	0.16	0.24	87.4	0.26	0.37	86.0
0.35	0.45	89.1	0.32	0.39	82.1	0.17	0.25	88.0	0.28	0.39	87.3

**Fig. 7** Performance curves of unimodal biometrics



**Table 4** Performance rates of public dataset samples measured at 0.1 % FAR

Feature type	Database	No. of Samples	Accuracy (%)
Iris	IR-TestV1	2000	92
Ear	IIT-Delhi	421	88
Palm vein	Dr. Badawi dataset	1000	88.2
Face	AR faces	1200	87

computing the total score from the similarity scores obtained in each modality. The person with fused score less than the threshold is determined as authenticated. This score level fusion is easier to implement than other fusion methods and doesn't involve any feature vectors directly. Hence the information loss is zero when compared to other fusion methods. This helps in reducing the false rate, in turn increasing the matching rate to greater than 94 % for any combination of traits. This is an improved performance when compared to unimodal system. This system is also tested with occluded samples such as reflections, contact lens and spectacles on iris, 3/4th close of eyelids, bunch of hairs covering the ear, poor visibility of palm veins and samples captured under different lighting conditions. The proposed system is responding to most of the occluded samples except iris sample with 3/4th close of eyelid, side faces and poor visibility of palm vein. Fusion gives a good performance even for occluded samples. These performances are not liable to compare with any literature works, since the features are acquired from the same person. No such public dataset with our chosen features. Hence we have tested our system with various combinations within our dataset and with the unimodal performances.

## 5 Conclusion

The focus of this work is to develop a versatile biometric system using cost effective acquisition devices and to emphasize the role of fusion in providing secure authentication when multiple features of an individual are handled. This work helps to learn the essentials in adapting multibiometrics for various practical applications based on the requirement. From the experiments, it is evident that the multibiometric system with fusion of different contactless biometric features improves the recognition rate than unimodal systems resulting in higher accuracy rate of 96.3 % with lower false rates. From the experimental result tables, it is shown clearly that every biometric feature plays its role excellently in the process of recognition and even when occlusions are involved. In the future, we would like to test the system for a very large dataset and focus especially in handling various occlusions and make it public for research works very soon.

**Acknowledgments** Glad to thank the persons who contributed their feature images to build this versatile multimodal dataset with great cooperation.

## References

1. Jain, A.K., Ross, A., Prabhakar, S.: An introduction to biometric recognition. *Circ. Syst. Video Technol. IEEE Trans. on* **14**(1), 4–20 (2004)
2. Bigun, J., Fiérrez-Aguilar, J., Ortega-García, J., Gonzalez-Rodriguez, J.: Combining biometric evidence for person authentication. In: *Advanced Studies in Biometrics*, pp. 1–18, Springer, Berlin Heidelberg (2005)
3. Ross, A., Jain, A.: Information fusion in biometrics. *Pattern Recogn. Lett.* **24**(13), 2115–2125 (2003)
4. Poornima, S., Nasreen, F., Prakash, A.D.S., Raghuraman, A.: Versatile and economical acquisition setup for dorsal palm vein authentication. *Procedia Comput. Sci.* **50**, 323–328 (2015)
5. Poornima, S., Subramanian, S.: Unconstrained iris authentication through fusion of RGB channel information. *Int. J. Pattern Recogn. Artif. Intell.* **28**(5), 1–18, 1456010 (2014)
6. Proença, H., Alexandre, L.: Iris recognition: an analysis of the aliasing problem in the iris normalization stage. *Int. Conf. Comput. Intell. Secur.* **2**, 1771–1774 (2006)
7. Daugman, J.: How iris recognition works. *IEEE Trans. Circ. Syst. Video Technol.* **14**(1), 21–30 (2004)
8. Masek, L.: Recognition of human iris patterns for biometric identification. *Doctoral Dissertation, Master's Thesis, University of Western Australia* (2003)
9. Wang, Q., Zhang, X., Li, M., Dong, X., Zhou, Q., Yin, Y.: AdaBoost and multi-orientation 2D gabor-based noisy iris recognition. *Pattern Recogn. Lett.* **33**(8), 978–983 (2012)
10. Viola, P., Jones, M.J.: Robust real-time face detection. *Int. J. Comput. Vis.* **57**(2), 137–154 (2004)
11. Kumar, A., Wu, C.: Automated human identification using ear imaging. *Pattern Recogn.* **45**, 956–968 (2012)
12. Guo, Z., Zhang, L., Zhang, D., Mou, X.: Hierarchical multiscale LBP for face and palmprint recognition. In *17th IEEE International Conference on Image Processing (ICIP)*, pp. 4521–4524 (2010)
13. Poornima, S., Subramanian, S.: Personal authentication through dorsal palm vein patterns. *Int. J. Appl. Eng. Res.* **10**(34), 27286–27290 (2015)