



An accurate multi-modal biometric identification system for person identification via fusion of face and finger print

Sidra Aleem¹ · Po Yang² · Saleha Masood¹ · Ping Li³ · Bin Sheng¹ 

Received: 27 October 2018 / Revised: 24 April 2019 / Accepted: 23 May 2019 /
Published online: 21 June 2019
© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Internet of things (IoT) have entirely revolutionized the industry. However, the cyber-security of IoT enabled cyber-physical systems is still one of the main challenges. The success of cyber-physical system is highly reliant on its capability to withstand cyberattacks. Biometric identification is the key factor responsible for the provision of secure cyber-physical system. The conventional unimodal biometric systems do not have the potential to provide the required level of security for cyber-physical system. The unimodal biometric systems are affected by a variety of issues like noisy sensor data, non-universality, susceptibility to forgery and lack of invariant representation. To overcome these issues and to provide higher-security enabled cyber-physical systems, the combination of different biometric modalities is required. To ensure a secure cyber-physical system, a novel multi-modal biometric system based on face and finger print is proposed in this work. Finger print matching is performed using alignment-based elastic algorithm. For the improved facial feature extraction, extended local binary patterns (ELBP) are used. For the effective dimensionality reduction of extracted ELBP feature space, local non-negative matrix factorization is used. Score level fusion is performed for the fusion. Experimental evaluation is done on FVC 2000 DB1, FVC 2000 DB2, ORL (AT&T) and YALE databases. The proposed method achieved a high recognition accuracy of 99.59%.

Keywords Face recognition · Finger print recognition · Cyber-physical systems · Extended local binary patterns · Local non-matrix factorization

This article belongs to the Topical Collection: *Special Issue on Smart Computing and Cyber Technology for Cyberization*
Guest Editors: Xiaokang Zhou, Flavia C. Delicato, Kevin Wang, and Runhe Huang

✉ Bin Sheng
shengbin@sjtu.edu.cn

Extended author information available on the last page of the article.

1 Introduction

Cyber-physical systems (CPS) have recently drawn a significant attention. The recent advances in technology have further led emphasis on the need of CPS. Cyber security is the critical challenge in the IoT enabled CPS [57, 58]. Cyber security is threatened by cyber attacks. Thus, for the successful deployment of secure CPS, cyber security is a critical factor. Biometric identification systems are one of the key identification systems in CPS [8, 14, 31, 59]. The success of CPS is highly reliant on its capability to withstand the cyber attacks. To overcome cyber attacks and provision of highly secure CPS, fusion of different biometric modalities is the need of hour. The main challenge in the design of highly accurate and secure deployment of biometric systems is the unavailability of high-quality data to ensure consistently high recognition results. Lighting conditions, distance to surveillance camera, equipment and operator training can significantly affect the performance. These limitations of unimodal biometrics can be alleviated by multi-modal biometrics to provide reliable authentication system [27, 48].

The consolidation of multiple biometric modalities can result in a more accurate identification system [11, 41, 42, 51]. Accuracy enhancement is the core motivation behind multi-biometric systems [17]. Accuracy of multi biometric systems is alleviated due to two reasons. Firstly, it reduces the overlap between the intra class distribution and effectively increases the dimensionality of the feature space [47]. Discriminatory information of a single trait can be affected by various reasons like noise, imprecision, and inherent drift due to aging [46]. Along with accuracy multi-modal biometrics have other advantages over unimodal like alleviation of the non-universality problem and reduction of enrollment errors, high degree of flexibility for user authentication and increased resistant to forgery [33]. Multi-modal biometrics are implemented by the fusion of unimodal biometrics. For the efficient performance of multi-modal biometric system an effective biometric fusion approach is required for the fusion of the information from multiple sources [22].

In this paper, the face and finger print biometric traits are fused to construct a multi-modal biometric recognition system. These two modalities are selected as they are most natural and acceptable means of biometrics. The elastic alignment algorithm is used to extract features from finger image. While ELBP is used for facial features extraction. Fusion of the chosen modalities is performed at the matching score level. The success of face recognition is highly reliant on the representation of high-dimensional face space. An efficient way is to find intrinsically low-dimensional subspace that effectively represent the underlying data. Local non-negative matrix factorization (LNMF) for dimensionality reduction of face subspace. LNMF not only achieves the dimensionality reduction by part based representation, but also focuses to preserve the local information, adds sparsity, retain only the most significant basis and provides more true parts based representation. As the face space have reminiscent features like lips, eyes, and nose, etc. [35–37]. So, LNMF completely suits the face feature space. The main motive of LNMF is to find reduced subspace without using class information. The non-negativity constraint leads to a sparsely distributed data coding that might be useful for extracting parts-based representation of data patterns with low feature dimensionality [52]. The major contributions of this paper are as follows:

- **Effective feature extraction** ELBP is used for facial feature extraction. In contrast to conventional local binary pattern(LBP) that only encodes the local information, ELBP also encodes the additional information that results in more discriminatory features. It provides more effective feature extraction.
- **Improved dimensionality reduction** LNMF is used for dimensionality reduction of face space. In contrast to traditional NMF it not only imparts non-negativity constraint,

but also highlights and make localized features more clear. As a result in contrast to NMF, LNMF provides more true parts-based representation.

2 Related work

A lot of work has been done in the field of multi-modal biometrics. Brunelli and Falavigna proposed a multi-modal biometrics based on the fusion of voice and face. Fusion is performed using weighted geometric average and for normalization hyperbolic tangent is used [9]. Kittler et al. used a variety of fusion techniques for voice and face. The experimentation was done on several fusion techniques including sum, product, minimum, median, and maximum rules and they have found that the sum rule outperformed others. It was found out that the sum rule is not significantly affected by the probability estimation errors and found out to be superior [32]. Multi-modal identification based on face and finger was proposed by Hong and Jain. First the database was pruned via face matching and after that finger print matching was performed [23]. Ben-Yacoub et al. developed a multi-modal biometric system based on face and voice. Several fusion techniques were analyzed including support vector machines, tree classifiers, and multilayer perceptrons. Of all these techniques Bayes classifiers were found to be the optimal method [6]. Ross and Jain developed multi-modal system by fusion of face, finger print, and hand geometry. Several fusion techniques were evaluated including sum, decision tree, and linear discriminant-based methods. Sum rule method outperformed the other paralleled techniques [43]. Komal [50] proposed iris and finger print fused multi-modal system. The proposed method used minutiae points and ridge to reduce spoofing attacks. The proposed system minimized the spoofing attack because of the acquisition source nature used for both modalities. Gaussian Mixture Modal (GMM) based Expectation Maximization (EM) algorithm was used for multi-modal biometric system [49]. Lu et al. used a combination of different classifiers for face recognition. The fusion is performed on match score level.

Chen and Te Chu et al. [12] proposed a biometric system with the combination of face and iris. The method made use of combination of ORL and CASIA databased for face and iris respectively. Besbes et al. [7] proposed a prototype multi biometric system with the integration of finger print and face. The method is proposed to overcome several limitations faced by the face recognition and finger print verification systems as a uni-models. The integration of these biometric traits resulted in an acceptable response time. Liao and Isa [38] introduced a multi biometric system, which combined face and iris for person identification. The fusion is performed at matching score level and enhanced support vector machine. An enhanced feature selection method was used for the feature extraction of face and iris. Al-khassaweneh et al. [5] fused finger print and iris for a person identification system. Haar wavelet decomposition is used for iris recognition and Euclidean distance from the bifurcation point to the core point is used for the finger print recognition. Shruthi et al. [45] performed fusion of finger and finger vein based on a new score level fusion approach. The performed fusion was a holistic and nonlinear.

3 Proposed methodology

The proposed biometric recognition system is based on the integration of two biometric traits i.e. finger print and face. It comprises of two main modules: Enrollment and Identification. The enrollment module is further divided into two sub-modules: finger print

enrollment and face enrollment. Finger print enrollment comprises of:- Pre-Processing: for artifacts exclusion, Center point detection: locating center point, Tessellation: region of interest is tessellated around the center point, Gabor filtering: enhancing the finger print, Feature extraction: the digital representation is further processed by a feature extractor to generate a compact but expressive representation, which are then stored in the database. Face enrollment:- Pre-processing: for elimination of unwanted information, Feature extraction: ELBP is used for the extraction of more discriminative features as compared to the conventional methods. Dimensionality reduction: ELBP extracted features are subjected to dimensionality reduction using LNMF for better representation as compared to the other parallel frames. The dimensionality reduced feature vectors are then stored in the database. Identification module comprises of:- Matching: fusion of individual scores obtained by each of the modality is performed. The final score is generated by the sum rule. Fused score is then compared with a pre-defined threshold. The final decision is made on the basis of thresholding. The overview of the proposed method is shown in Figure 1.

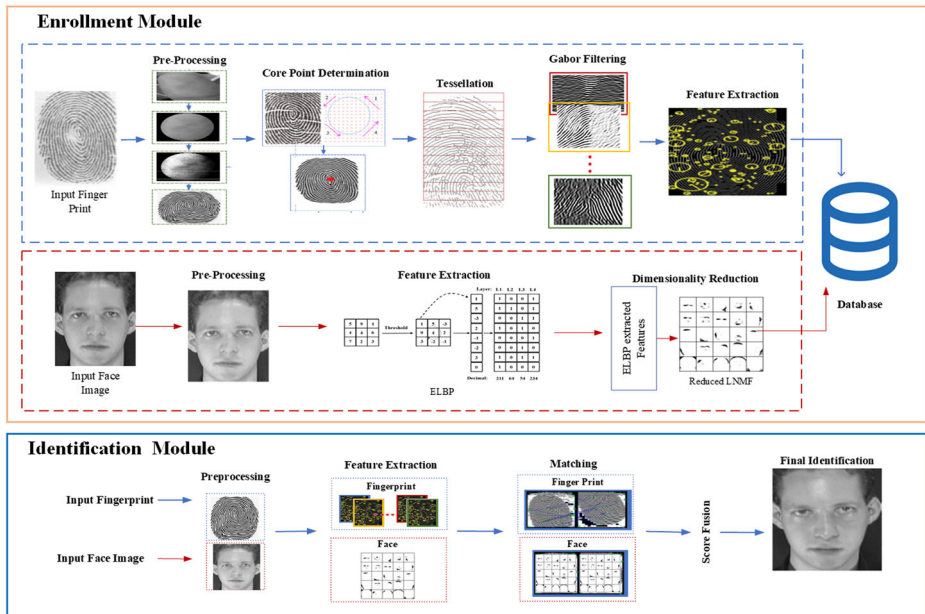


Figure 1 System Overview: The system consists of two main modules. **a** Enrollment Module:- Finger print: Pre-Processing: Artifacts are removed from the input images. Core point determination: Orientation field masking is used for core point determination. Tessellation: Region of interest around the the detected core point is tessellated. Gabor Filtering: Images are subjected to gabor filtering to make them appropriate for feature extraction. Feature Extraction: Features are extracted from Gabor filtered images. The resulting feature vectors are then stored in the database. Face: Pre-Processing: Artifacts are removed from the input images. Feature Extraction: Features are extracted using enhanced local binary pattern (ELBP). Dimensionality reduction: The extracted feature vectors dimensionality is reduced using local non-matrix factorization(LNMF). The reduced feature vectors are then stored in the database. Identification Module:- Matching: Extracted features are matched with the feature vectors stored in the database. Score Fusion: Scores obtained from both the modalities are fused. On the basis of score final identification is performed

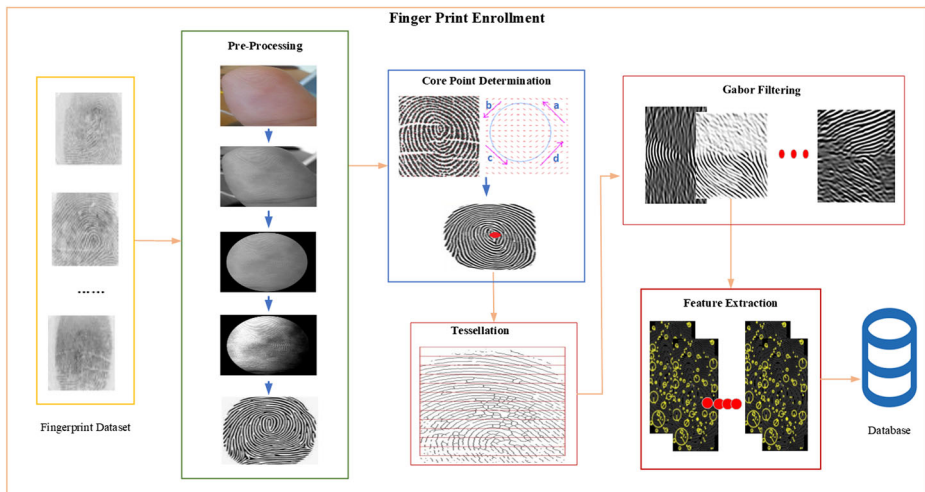


Figure 2 Finger print Enrollment Module:- Pre-Processing: Finger prints are subjected to preprocessing for removing artifacts. Core point detection: Core point is detected for alignment. Tessellation: The region of interest around the detected core point is tessellated. Feature Extraction: Tessellated images are subjected to Gabor filtering in six different directions. The feature are then extracted from Gabor Filtered images. The resulting feature vectors are stored in the database

3.1 Finger print recognition

The individuality of finger prints have been known since decades. They are utilized widely for biometric identification due to their invariance and singularity. The main factor responsible for widespread use of finger prints is the numerous advantages they offer like matching speed and accuracy [29]. A series of operations are to be performed to make them appropriate for the accurate feature extraction [18]. Despite the discriminatory nature of finger prints, the development of a reliable finger print matching system is still a very challenging task. The overview of the proposed finger print module is given in Figure 2.

3.1.1 Pre-processing

The quality of finger print image is adversely affected by the image acquisition modality. As a result, breaks between different features points occur that later on affect the identification process. Furthermore, there is a poor contrast between foreground and highly noisy background. All these artifacts must be removed before the identification process. Thus pre-processing is vital to remove all these artifacts.

3.1.2 Center point detection

To make a robust and accurate identification system it is necessary to make finger prints invariant to rotation, translation and scale. Core point is used as a reference point in this work. For core point detection, orientation field mask is used. The chosen method exploits the fact that core point have specific pattern orientation field. These patterns appear as loops

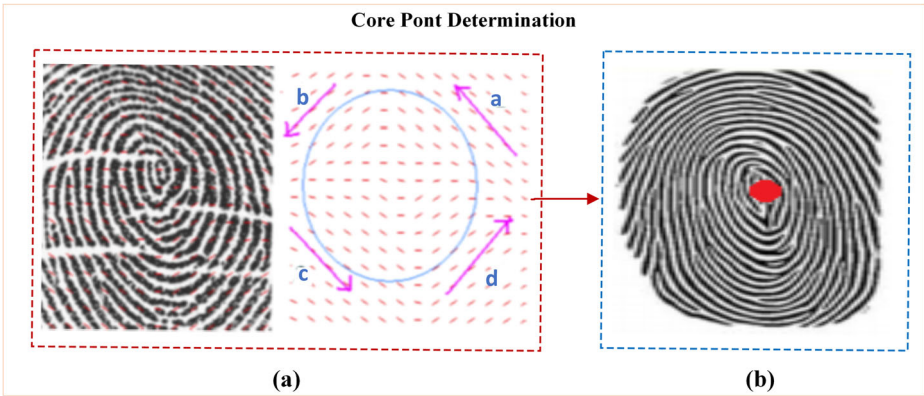


Figure 3 Core Point Detection: **a** Determination of core point using orientation field mask. **b** Detected core point

in the region containing core points as shown in Figure 3. The mask gives a maximum magnitude of convolution in the region of core point. The angle in the regions a, c ranges from 40 to 90 degrees. While the angle in the regions b, c ranges from 0 to 45. Whenever such pattern exists in the finger print, it can be regarded as a core point. We empirically generated a mask to locate and detect similar pattern in the orientation field shown in Figure 3a. The orientation field mask is shown below in Figure 3a. The response achieved by applying orientation mask corresponds to the nature of orientation field around that point. Loop field strength (*LFS*) in the orientation field is given by:

$$LFS = \sum_{x=-s/2}^{x=s/2} \sum_{y=-s/2}^{y=s/2} IM[X, Y]O[i + x, j + y] \tag{1}$$

where *s* is mask size, *I* is the image and *M* is the mask. *LFS* is calculated for each element in the orientation map. After that threshold is applied to the obtained magnitude array.

3.1.3 Tessellation of region of interest

After the successful detection of center point, the next step is to tessellate the region of interest around the core point. The region of interest corresponds to all the sectors around the core point. We have divided the region of interest into 80 sectors. Let *I*(*x*, *y*) denote the gray level at pixel in an finger print image and let *I*(*x_c*, *y_c*) denote the reference point and *S_i* as sector.

$$S_i = (x, y) | b(T_i + 1) \leq r < b(T_i + 2), \theta_i \leq \theta < \theta_{i+1}, 1 \leq x \leq N, 1 \leq y \leq M \tag{2}$$

where $T_i = i \text{ div } k$, $\theta_i = (i \text{ mod } k) \times (2\pi/k)$, $r = \sqrt{(x - x_c)^2 + (y - y_c)^2}$ and $\theta = \tan^{-1}((y - y_c)/(x - x_c))$.

3.1.4 Gabor filtering

Feature vector in this work is composed of the ordered calculation of features of all the sectors specified by tessellation. As a result, feature vectors have local feature information. The ordered calculation of features from all the sectors helps to attain global feature information.

Thus, both local and global information is retained by our method. The local information is captured by feature elements and ordered calculation from the tessellation achieves the global information among the obtained local features. After having the information from local features, the next step is to decompose the discriminatory information achieved by each of the sectors independently. Gabor filters are well known for capturing information in separate bandpass channels and are capable of converting this information in terms of spatial frequency by decomposing the information into biorthogonal components [30]. The Gabor filter in spatial domain is given by:

$$G(x, y, f, \theta) = \exp \left\{ -\frac{1}{2} \left[\frac{x'^2}{\delta^2} + \frac{y'^2}{\delta^2} \right] \right\} \tag{3}$$

$$x' = x \sin\theta + y \cos\theta \tag{4}$$

$$y' = x \cos\theta - y \sin\theta \tag{5}$$

where f is the frequency of the sinusoidal plane wave along the direction θ from the x -axis, $\delta_{x'}$ and $\delta_{y'}$ are the space constants of the Gaussian envelope along x' and y' axes, respectively.

The region of interest is obtained by filtering finger print with Gabor filter in six different directions. These directions are mandatory to capture both local and global information of the finger print. Filtering with Gabor filter is performed in spatial domain with mask of size 30x30. The convolution with Gabor filtering is the chief contributor to overall execution time for feature extraction.

3.1.5 Feature vector formation

After filtering, the finger print is appropriate for feature extraction. Absolute deviation from the mean of Gabor filtered image for every individual sector is calculated. After that average of the mean from all the sectors obtained as a result of tessellation is calculated. This average absolute deviation is used as a feature to be stored in feature vector. Feature vector is formulated as:

$$V_{i\theta} = \frac{1}{2} \left(\sum_{n_i} |F_{i\theta}(x, y) - P_{i\theta}| \right) \tag{6}$$

where $f_{I\theta}$ is the direction filtered image for the tessellated sectors S_i , $\theta \in \{0, \frac{\pi}{6}, \frac{\pi}{3}, \frac{\pi}{2}, \frac{2\pi}{3}, \frac{5\pi}{6}\}$, the feature value $V_{i\theta}$ is the average absolute deviation from the mean, n_i is the number of pixels in and $P_{i\theta}$ is the mean of pixel values of in sector S_j .

3.2 Face recognition

3.2.1 Pre-processing

One of the main challenges of face recognition is the illumination variance. Images having uncontrolled lighting conditions suffer from non-uniform contrast because of the uneven distribution of the gray levels. In order to equally distribute the intensity level and to increase the contrast of the input image, histogram equalization [16] is used for the preprocessing. Figure 4 shows the proposed face recognition system.

$$h = \frac{CDF(V) - CDFmin}{I - CDFmin} (L - 1) \tag{7}$$

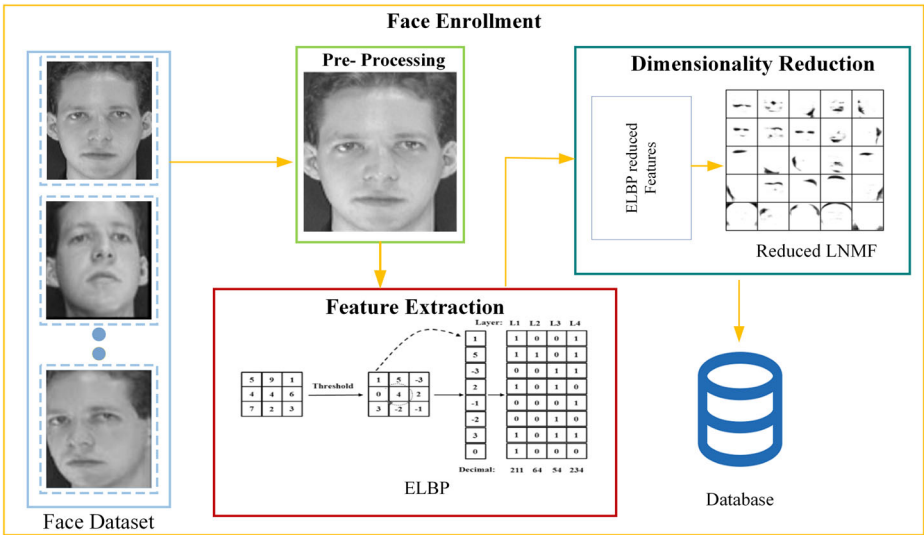


Figure 4 Face Enrollment Module:- Pre-processing: Removed the artifacts. Feature Extraction: Features are extracted using Extended Local Binary Pattern. Dimensionality reduction: Extracted features are subjected to dimensionality reduction using LNMF. The reduced dimensionality templates are then stored in database

where $I(x, y)$ is the input image, M and N corresponds to the total number of rows and columns, CDF is the cumulative distribution function, V is the pixel value and L refers to the gray levels.

3.2.2 Feature extraction

LBP has been extensively used for feature extraction in face recognition. It is a pixel based texture extraction method that has achieved remarkable performance along with low computation cost [24]. The key issue affecting the performance of local appearance based face recognition is the method of finding the most discriminative facial areas [53]. Recently a lot of variants of LBP have been developed for the improvement of discriminative capability of LBP. To further enhance the discriminative power of LBP, additional information about the local patterns and structures can be encoded [4]. To enhance the discriminative power of LBP, the ELBP is used in this work [25, 26].

In addition to the comparison between the central and its neighbouring pixels done by the conventional LBP, ELBP also encodes the exact gray level values using some additional binary units. ELBP consists of several layers. These layers have various LBP codes that encode the gray level differences. ELBP first computes the LBP code. The LBP is calculated by thresholding the 3x3 neighbourhood of all the pixels in image to the central pixel of the window. The LBP labels computed over this region are used as texture descriptors [21]. Formally, given a pixel at (x_c, y_c) , the resulting LBP can be expressed in decimal form as:

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(i_p - i_c)2^p, \tag{8}$$

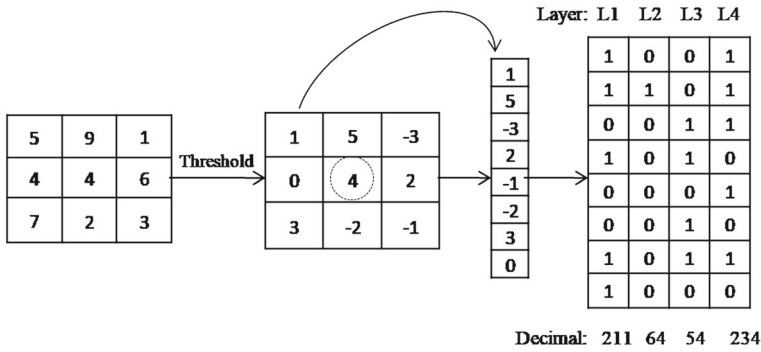


Figure 5 Extended Local Binary Pattern (ELBP) work flow. On first step the conventional LBP is extracted. The pixels in 3x3 window are compared with the central pixel. The pixel values that are less than central pixel gets the code of zero and the ones that are greater than central pixels get 1. L1 has the code of sign of gray level values difference in LBP. 0 is placed in L1 where the difference is negative and 1 is placed where the difference is positive. In the successive layers the actual code for gray level value is stored

where i_c and i_P are the gray-level values of the central pixel and P are the surrounding pixels in the circle neighborhood with a radius R , and function $s(x)$ is defined as:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \tag{9}$$

The working mechanism of ELBP is shown in Figure 5. After calculating LBP, the ELBP stores the code of the absolute sign of gray level difference in the first layer (L1). For a positive gray level difference, value 1 and for negative gray level difference, value 0 is stored at L1 respectively. In the successive layers the actual gray level value is encoded. So, first the absolute sign of gray level difference is encoded on L1 and then the gray level values in their corresponding binary representations are stored in the following layers [19]. Consequently, LBP does not yield discriminative features for the similar texture features. In contrast, ELBP encodes the additional information in layers that boosts up its discriminative power.

3.2.3 Subspace learning

The high-dimensional feature vectors contain the redundant representation of face space. To achieve more compact face space, feature space can be reduced [54]. This reduction becomes even more crucial for face recognition system for real-time systems. NMF is used widely for dimensionality reduction. However, there are two main drawbacks that affect the dimensionality reduction through NMF. First is the high computational cost associated with the high dimension matrix reduction and other is repetitive learning that needs to be performed with the class and training samples update [13]. Subspace learning performs the dimensionality reduction by mapping data set from high to low dimensionality [44]. To overcome these issues, we reduced the dimensionality of the feature space by LNMF. LNMF ensures three factors: a) Sparsity: the motive is to make weight coefficient matrix as sparse as possible. b) Orthogonal basis: ensure that basis are orthogonal by making the

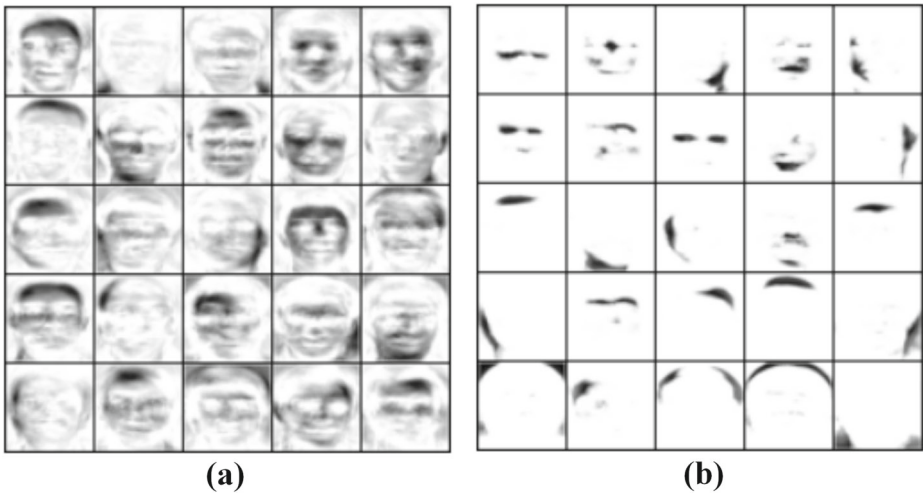


Figure 6 Comparison of NMF and LNMF. **a** The dimensionality reduction achieved by NMF. The basis have a lot of redundant information. **b** The dimensionality reduction achieved by LNMF. The representations are both part based and localized. LNMF has also eliminated the redundant information much efficiently

difference between base of minimum redundancy. **c**) Retain useful information: retain only the most significant information of the base. LNMF cost function is given as:

$$F = \sum_{i,j} (V_{ij} \log \frac{V_{ij}}{WH_{ij}} - V_{ij} + WH_{ij}) - \alpha \sum_{i,j} p_{ij} - b \sum +iq_{ij} \tag{10}$$

where the matrix X is decomposed in to W and H . W refers to the base image i.e. the feature space and $H = [h_1, h_2, \dots, h_3]$ for weight coefficient. ELBP and LNMF have two main advantages. **a**) the effective face texture information extraction, **b**) dimensionality reduction of the facial subspace. These two factors effectively improve the recognition accuracy. As evident from Figure 6, LNMF learns such basis components which not only lead to non-subtractive representations, but also highlights and make localized features much clear. As a result in contrast to NMF, LNMF is more true parts-based representations. Depending on the number of components and dimension.

4 Identification module

4.1 Finger print matching

An alignment-based elastic matching algorithm [28] is used for finger print matching. This algorithm finds the correspondence among the feature points without resorting to a thorough exploration. The algorithm first performs alignment of the feature points between the query image template and the templates stored in the database. The alignment is done by estimating parameters through transformations including translation, scaling and rotation.

$$A = ((x_1^A, y_1^A, \theta_1^A)^T, \dots, (x_a^A, y_a^A, \theta_a^A)^T) \tag{11}$$

$$B = ((x_1^B, y_1^B, \theta_1^B)^T, \dots, (x_b^B, y_b^B, \theta_b^B)^T) \tag{12}$$

where A is the query image template, B are the templates stored in the database, a are the feature points in A and b are the feature points in B respectively.

The aligned feature vectors are then converted to strings in the polar coordinate system. The conversion represents each of the pattern as two symbolic strings. The strings are formed by concatenating every feature point in an increasing order of radial angles as:

$$A^* = ((r_1^A, e_1^A, \theta_1^A)^T, \dots, (r_a^A, e_a^A, \theta_a^A)^T) \tag{13}$$

and

$$B^* = ((r_1^B, e_1^B, \theta_1^B)^T, \dots, (r_b^B, e_b^B, \theta_b^B)^T), \tag{14}$$

where r_i^A is the radius, e_i^A is the radial angle and θ_i^A is orientation of normalized feature vector with respect to reference feature (x, y, θ) .

After the conversion, matching between the converted string A^* and B^* is performed using a modified dynamic programming algorithm. The algorithm performs the similarity analysis by computing the the distance A^* and B^* . The minimum distance between A^* and B^* is used to establish the correspondence between the A^* and B^* feature vectors. The matching score was calculated as:

$$S = \frac{100M_{AB}^2}{ab} \tag{15}$$

where ab are the features points in the query image temple and templates stored in the database, M_{AB} are the feature points which lie in the bounding boxes of the template. The bounding box of a feature point defines the suitable acceptance in the locations of the corresponding feature of input with reference to the template.

4.2 Face matching

The query image I with dimension $M \times N$ is first converted into a vector x where $x \in R^{MN}$. Feature vector x represents the intensity at the pixel (m, n) as the $(mN + n)^{th}$ dimension in x . The similarity between the query image feature vector x and a feature vectors y stored in the database is evaluated on the basis of Euclidean distance.

$$d^2(x, y) = \sum_{k=1}^{MN} (x^k - y^k)^2 = (x - y)^T(x - y) \tag{16}$$

where d represents the calculated distance, $k = \{1, 2, \dots, MN\}$ is the index, M corresponds to total number of rows and N is total number of columns. Based on the similarity analysis, the feature vector y from the database having the minimum Euclidean distance with the input feature vector x is considered as a closest possible match.

4.3 Fusion

The scores obtained from the matching stage of both finger print and face do not lie in the same numerical range. So, the scores retrieved from both the modalities first need to be transformed to the same numerical range. The scores normalization is done by min-max approach [2].

$$w_{ik}^* = \frac{w_{ik} - \min^f(w_i)}{\max^f(v_i) - \min^f(v_i)} \tag{17}$$

where w_{ik} are the un-normalized scores from the matching stage, w_{ik}^* is the normalized score, \min and \max represents the minimum and maximum scores of the feature to be normalized and f corresponds to the modality under consideration. The sum-rule based

fusion is then used for the fusion of normalized scores of the both modalities. After a set of normalized scores (x_1, x_2, \dots, x_m) is obtained for both finger and face fusion is performed.

$$f_s = w_1x_1 + \dots + w_mx_m \quad (18)$$

where f_s is the final score after fusion, w is the weight associated to each individual score. The fused score f_s is then compared to a pre-defined threshold T . If $f_s \geq T$, then the individual is declared as the match, otherwise declared as an imposter.

5 Experimental results

5.1 Datasets

FVC 2000 DB1 [10] and FVC 2000 DB2 [39] are used for the evaluation on finger print. The samples in FVC 2000 DB1 [10] are taken over the different time interval, using low-cost optical sensor, and the resolution of images is 300 x 300. In FVC 2000 DB2 [39] there are four samples per subject, low-cost capacitive sensor is used for acquisition, and the resolution of the images is 256 x 364. ORL (AT&t) and YALE [15] databases are used for performance evaluation on facial images. ORL is standard AT&T database that consists of 400 images of 40 subjects with 10 samples per subject. The images are of resolution 92 x 112. The Yale Face Database [15] consists of 15 subjects with 11 samples per subject.

5.2 Recognition accuracy based comparison

To evaluate the performance of proposed method, it is compared with other existing multi-modal and unimodal biometric system. Following experiments have been carried out for the comparative analysis. Experiment1: The unimodal biometrics (i.e. finger print and face) which are used for the proposed multi-modal method are compared with the proposed method. For finger print evaluation FVC 2000 DB1 [10] and DB2 [39] and for face ORL (AT&T) and YALE [15] databases are used. The statistical details of this experiment are given in Table 1. Recognition accuracy is the chosen evaluation parameter for the performance analysis of this experiment. Recognition accuracy of different modalities is compared. From Table 1. it is evident that proposed multi-modal system achieved highest accuracy in contrast to the unimodal systems. The proposed method achieved 99.59% accuracy. Comparison of the unimodal versus multi-modal system can be analyzed in Figure 7.

Table 1 Accuracy comparison between unimodal and proposed multi-modal system

Modality	Database	Recognition accuracy (%)
Finger print	FVC 2000 DB1 [10]	98.00
Finger print	FVC 2000 DB2 [39]	98.30
Face	ORL (AT&T)	97.20
Face	YALE [15]	97.71
Proposed	FVC 2000 DB1 , DB2, ORL, YALE	99.59

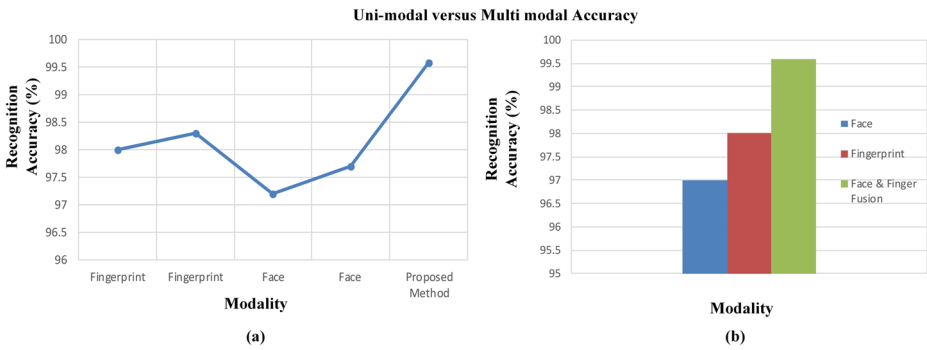


Figure 7 Accuracy comparison of unimodal versus multi-modal biometrics: **a** Recognition accuracy comparison between of the proposed method and unimodal biometrics. The proposed method achieved the highest recognition accuracy. **b** Recognition accuracy comparison of the proposed method and the modalities fused for our multi-modal system i.e. finger and face. The proposed multi-modal clearly surpass the two unimodalities

According to the performed experiment, the proposed method outperforms the unimodal systems. The unimodal system based on face yielded 97% accuracy, finger print based identification system resulted in 98% accuracy, whereas the proposed method achieved 99.59% recognition accuracy. Experiment 2: The proposed multi-modal recognition system is compared with other existing multi-modal techniques. Various multi modalities including face + palm print [20], side + gait [56], hand geometry + palm print [34], palm print+middle finger [55], finger print+ iris [1], iris + palm print [3] have been compared with our proposed combination. Our multi-modal biometric system achieves better recognition accuracy in contrast to other existing techniques.

Table 2 gives the statistical details of this comparison. From the experimental results, it is evident that fusion of face and finger print achieved highest recognition accuracy in contrast to other modalities fusion. So, the natural acquisition of both modalities and better accuracy gives it a lead over other paralleled methods. Accuracy comparison of the proposed method with other methods is shown in Figure 8. Based on the performed comparison, it can be analyzed the recognition accuracy for face and palm print [20] is 81.46%, side face and gait [56] is 93.30%, hand geometry and palm print [34] is 94.59%, palm print and middle finger [55] is 92.67%, for face and finger print [40] is 92.67%, for finger print and iris is [1] is

Table 2 Recognition accuracy comparison of various multi-modal methods

Method	Modalities	Recognition accuracy (%)
Guo et al. [20]	Face and palm print	81.46
Zhao and Bhanu et al. [56]	Side face and gait	93.30
Kumar et al. [34]	Hand geometry and palm print	94.59
Zhang et al. [55]	Palm print and middle finger	92.67
Korves et al. [40]	Face and finger print	92.67
Abdolahi et al. [1]	Finger print and Iris	98.20
Aguilar et al. [3]	Iris and palm print	98.20
Our Proposed	Face and finger print	99.59

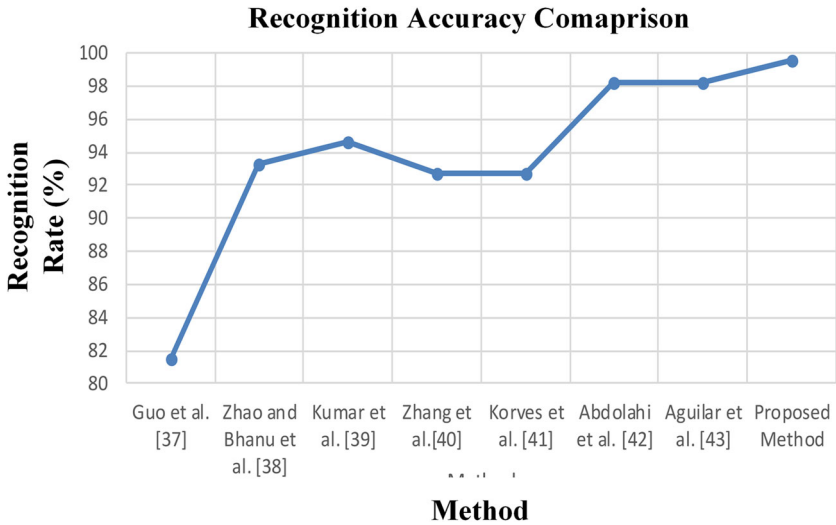


Figure 8 Recognition accuracy comparison of proposed method and other multi-modal approaches. The proposed method achieved highest recognition accuracy as compared to all other techniques. Hence our method is superior as compared to other paralleled frame works

98.20%, iris and palm print [3] is 98.20%. In contrast to all these techniques, the proposed method achieved the highest recognition accuracy of 99.59%.

5.3 Results in terms of error rate

Error rate is the evaluation metric that decides the performance of the biometric system. Equal error rate (EER) corresponds to that value of error where both False acceptance rate(FAR) and False rejection rate(FRR) are equal. The lower the value of EER better is the performance of the biometric system. To evaluate the performance of the proposed fused modalities with other existing multi-modal systems their corresponding EER have been compared. The statistical details of this comparison are given in Table 3. The proposed multi-modal system has the least EER as compared to all the other paralleled multi-modal systems. Figure 9a shows the EER with respect to FAR and FRR for the proposed method and other existing techniques. With proposed method EER is 0.035, with face EER is 0.0528 and for finger EER is 0.042. Hence the proposed method achieved lower EER in contrast to paralleled modalities.

Table 3 Equal error rate comparison of different modalities

Modality	Database	Equal error rate (%)
Finger print	FVC 2000 DB1 [10]	0.40
Finger print	FVC 2000 DB2 [39]	0.45
Face	ORL (AT&T)	0.5
Face	YALE [15]	0.83
Proposed	FVC 2000 DB1 , DB2, ORL, YALE	0.035

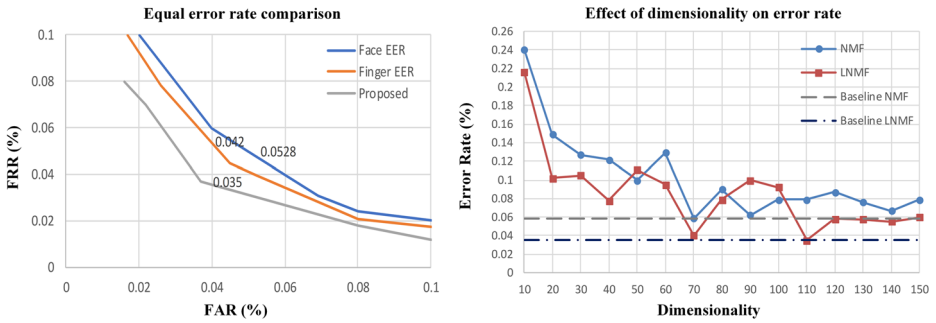


Figure 9 Error rate comparison. **a** The error rate of face, finger and the proposed method are compared. The proposed method achieved the lowest EER. **b** The effect of dimensionality reduction on the error rate. LNMF baseline shows that it achieved the lowest EER

Furthermore, we evaluated the EER with respect to dimensionality change. The effect of dimensionality change has been observed using NMF and LNMF and comparing the corresponding errors rates of the two techniques. The corresponding error rate of both methods clearly shows that LNMF have less error rate as compared to NMF. Error rate for LNMF ranges from 0.22 to 0.061 and for NMF it ranges from 0.24 to 0.09. The LNMF baseline shows that EER of 0.035 and NMF baseline shows that EER of 0.059 is achieved with LNMF and NMF respectively. These metrics clearly depict that our proposed method achieved the lowest EER as compared to the other existing techniques. In addition it uses LNMF with which lowest EER is achieved.

5.4 Subspace learning technique comparison

Subspace learning has a very significant impact on the recognition accuracy. As the opted subspace learning method decides which of the features are to be retained in the final feature vector. So, such a sub space learning technique should be opted that provides the optimal performance. After experimentation and comparison we found out that LNMF provides optimal solution for the proposed method. The impact of subspace is evaluated in terms of error rate and the basis components retained by the resulting basis. Though NMF has been widely used for subspace learning in face recognition. However the high computational cost and repetitive learning are two major drawbacks associated with NMF. A comparative analysis has been done to evaluate the performance of LNMF and NMF for the proposed method. The error rate of LNMF is less as compared to NMF as shown in Figure 9b. As error rate is one of the crucial factors that attributes to performance of biometric system. So, LNMF provides the optimal solution in contrast to NMF.

The dimensionality of chosen subspace learning method also impacts the performance as the basis components vary tremendously according to the chosen dimension. Figure 10 shows the visual impact of varying dimensionality on the basis learned correspondingly. The basis components are clearly depicting that LNMF retains more localized information. While NMF components have non useful information as well. With increasing dimensionality the learned basis provide more localized information. Figure 9 shows the impact of varying dimensionality on error rate. Dimensionality is varied over the range of 10 to 150. Error rate varies corresponding to dimensionality. For NMF the lowest EER is observed at dimensionality of 70 and for LNMF it is observed at 110.

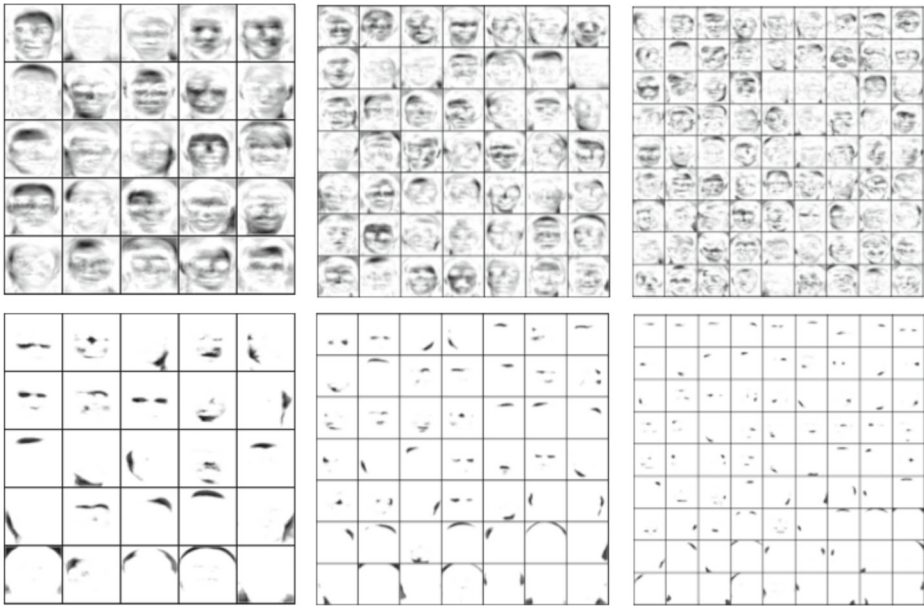


Figure 10 Impact of varying dimensionality on the basis components. Dimensions 25 (column 1), 49 (column 2) and 81 (column 3). The first row corresponds to dimensionality reduction achieved by NMF. The basis have a lot of redundant and unuseful information. The second row corresponds to dimensionality reduction achieved by LNMF. The representations are both part based and localized. LNMF has also eliminated the redundant information much efficiently

6 Conclusion

An accurate multi-modal system based on the fusion of face and finger is developed. In contrast to the other existing techniques, the fusion of the proposed modalities achieved a high recognition accuracy of 99.59%. Experimental result showed that lowest EER of 0.035% is achieved with the proposed method. For the extraction of feature with more discriminatory power ELBP is used. These enhanced discriminative power features resulted in more effective feature extraction and high recognition accuracy. LNMF is used for dimensionality reduction to overcome the decline in recognition that occur due to dimensionality reduction technique. In addition to non-negativity, LNMF make localized features more clear that in turn results in more true parts-based representation. In future, we aim to further improve the recognition accuracy to make the proposed method even more suitable for real time applications.

Acknowledgments This work was supported in part by the National Natural Science Foundation of China under Grant 61872241 and Grant 61572316, in part by the National Key Research and Development Program of China under Grant 2017YFE0104000 and Grant 2016YFC1300302, in part by the Macau Science and Technology Development Fund under Grant 0027/2018/A1, and in part by the Science and Technology Commission of Shanghai Municipality under Grant 18410750700, Grant 17411952600, and Grant 16DZ0501100.

References

1. Abdolahi, M., Mohamadi, M., Jafari, M.: Multimodal biometric system fusion using fingerprint and iris with fuzzy logic. *International Journal of Soft Computing and Engineering* **2**(6), 504–510 (2013)
2. Aboshosha, A., El Dahshan, K.A., Karam, E.A., Ebeid, E.A.: Score level fusion for fingerprint, iris and face biometrics. *Int. J. Comput. Appl.* **111**(4), 47–55 (2015)
3. Aguilar, G., Sánchez, G., Toscano, K., Nakano, M., Pérez, H.: Multimodal Biometric System Using Fingerprint. In: *ICIAS 2007. International Conference on Intelligent and Advanced Systems*, 2007, pp. 145–150. IEEE (2007)
4. Ahonen, T., Hadid, A., Pietikainen, M.: Face description with local binary patterns: Application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(12), 2037–2041 (2006)
5. Al-khassaweneh, M., Smeirat, B., Ali, T.B.: A hybrid system of iris and fingerprint recognition for security applications. In: *2012 IEEE Conference on Open Systems (ICOS)*, pp. 1–4. IEEE (2012)
6. Ben-Yacoub, S., Abdeljaoued, Y., Mayoraz, E.: Fusion of face and speech data for person identity verification. *IEEE Trans. Neural Netw.* **10**(5), 1065–1074 (1999)
7. Besbes, F., Trichili, H., Solaiman, B.: Multimodal Biometric System Based on Fingerprint Identification and Iris Recognition. In: *2008 3rd International Conference on Information and Communication Technologies: from Theory to Applications* (2008)
8. Biswas, D., Everson, L., Liu, M., Panwar, M., Verhoef, B.E., Patki, S., Kim, C.H., Acharyya, A., Van Hoof, C., Konijnenburg, M., Van Helleputte, N.: CorNET: Deep learning framework for PPG-based heart rate estimation and biometric identification in ambulant environment. *IEEE Trans. Biomed. Circuits Syst.* **13**(2), 282–291 (2019)
9. Brunelli, R., Falavigna, D.: Person identification using multiple cues. *IEEE transactions on pattern analysis and machine intelligence* **17**(10), 955–966 (1995)
10. Cappelli, R., Erol, A., Maio, D., Maltoni, D.: Synthetic fingerprint-image generation. In: *Proceedings. 15Th International Conference on Pattern Recognition*, 2000, Vol. 3, pp. 471–474. IEEE (2000)
11. Chaki, J., Dey, N., Shi, F., Sherratt, R.S.: Pattern mining approaches used in sensor-based biometric recognition: a review. *IEEE Sensors J* **19**(10), 3569–3580 (2019)
12. Chen, C.H., Te Chu, C.: Fusion of face and iris features for multimodal biometrics. In: *International Conference on Biometrics*, pp. 571–580. Springer (2006)
13. Chen, W.S., Pan, B., Fang, B., Li, M., Tang, J.: Incremental nonnegative matrix factorization for face recognition. *Mathematical problems in Engineering* **2008** (2008)
14. Cheng, K.H.M., Kumar, A.: Contactless biometric identification using 3D finger knuckle patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* pp 1–15 (2019)
15. Christiansen, E., Kwak, I.S., Belongie, S., Kriegman, D.: Face box shape and verification. In: *International Symposium on Visual Computing*, pp. 550–561. Springer (2013)
16. Dharavath, K., Talukdar, F.A., Laskar, R.H.: Improving face recognition rate with image preprocessing. *Indian J. Sci. Technol.* **7**(8), 1170–1175 (2014)
17. Fu, B., Yang, S.X., Li, J., Hu, D.: Multibiometric cryptosystem: Model structure and performance analysis. *IEEE Trans. Inf. Forensics Secur.* **4**(4), 867–882 (2009)
18. Gaensslen, R.E., Ramotowski, R., Lee, H.C.: *Advances in fingerprint technology*. CRC press (2001)
19. Guo, Z., Zhang, L., Zhang, D.: A completed modeling of local binary pattern operator for texture classification. *IEEE Trans. Image Process.* **19**(6), 1657–1663 (2010)
20. Guo, Z., Zhang, L., Zhang, D., Mou, X.: Hierarchical multiscale lbp for face and palmprint recognition. In: *2010 17Th IEEE International Conference on Image Processing (ICIP)*, pp. 4521–4524. IEEE (2010)
21. Hadid, A., Pietikainen, M., Ahonen, T.: A discriminative feature space for detecting and recognizing faces. In: *CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2004, Vol. 2, pp. II–II. IEEE (2004)
22. Heo, J., Kong, S.G., Abidi, B.R., Abidi, M.A.: Fusion of visual and thermal signatures with eyeglass removal for robust face recognition. In: *2004. CVPRW'04. Conference on Computer Vision and Pattern Recognition Workshop*, pp. 122–122. IEEE (2004)
23. Hong, L., Jain, A.: Integrating faces and fingerprints for personal identification. In: *Asian Conference on Computer Vision*, pp. 16–23. Springer (1998)
24. Huang, D., Shan, C., Ardabilian, M., Wang, Y., Chen, L.: Local binary patterns and its application to facial image analysis: a survey. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **41**(6), 765–781 (2011)
25. Huang, D., Wang, Y., Wang, Y.: A robust method for near infrared face recognition based on extended local binary pattern. In: *International Symposium on Visual Computing*, pp. 437–446. Springer (2007)

26. Huang, Y., Wang, Y., Tan, T.: Combining Statistics of Geometrical and Correlative Features for 3D Face Recognition. In: *BMVC*, pp. 879–888. Citeseer (2006)
27. Jagadiswary, D., Saraswady, D.: Biometric authentication using fused multimodal biometric. *Prog. Comput. Sci.* **85**, 109–116 (2016)
28. Jain, A., Hong, L.: On-line fingerprint verification. In: *Proceedings of the 13th International Conference on Pattern Recognition*, 1996, Vol. 3, pp. 596–600. IEEE (1996)
29. Jain, A.K., Bolle, R., Pankanti, S.: *Biometrics: personal identification in networked society*, vol. 479 Springer Science & Business Media (2006)
30. Jain, A.K., Prabhakar, S., Hong, L.: A multichannel approach to fingerprint classification. *IEEE Trans. Pattern Anal. Mach. Intell.* **21**(4), 348–359 (1999)
31. Khaitan, S.K., McCalley, J.D.: Design techniques and applications of cyberphysical systems: a survey. *IEEE Syst. J.* **9**(2), 350–365 (2015)
32. Kittler, J., Hatef, M., Duin, R.P., Matas, J.: On combining classifiers. *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(3), 226–239 (1998)
33. Kumar, A., Kanhangad, V., Zhang, D.: A new framework for adaptive multimodal biometrics management. *IEEE Trans. Inf. Forensics Secur.* **5**(1), 92–102 (2010)
34. Kumar, B.V.K.V., Mahalanobis, A., Juday, R.D.: *Correlation Pattern Recognition*. Cambridge University Press, Cambridge (2005)
35. Lee, D.D., Seung, H.S.: Learning the parts of objects by non-negative matrix factorization. *Nature* **401**(6755), 788 (1999)
36. Lee, D.D., Seung, H.S.: Algorithms for non-negative matrix factorization. In: *Advances in Neural Information Processing Systems*, pp. 556–562 (2001)
37. Li, S.Z., Hou, X.W., Zhang, H.J., Cheng, Q.S.: Learning spatially localized, parts-based representation. In: *CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2001, Vol. 1, pp. I–I. IEEE (2001)
38. Liao, H.F., Isa, D.: Feature selection for support vector machine-based face-iris multimodal biometric system. *Expert Systems with Applications* **38**(9), 11,105–11,111 (2011)
39. Maio, D., Maltoni, D., Cappelli, R., Wayman, J.L., Jain, A.K.: *Fvc2000: Fingerprint verification competition*. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(3), 402–412 (2002)
40. Nadel, L.D., Korves, H.J., Ulery, B.T., Masi, D.: Multi-biometric fusion-from research to operations. In: *Biometric Consortium 2005 Conference*, pp. 19–21. Citeseer (2005)
41. Nandakumar, K., Chen, Y., Dass, S.C., Jain, A.K.: Likelihood ratio-based biometric score fusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(2), 342–347 (2008)
42. Neal, T.J., Woodard, D.L.: You are not acting like yourself: A study on soft biometric classification, person identification, and mobile device use. *IEEE Transactions on Biometrics, Behavior, and Identity Science* **1**(2), 109–122 (2019)
43. Ross, A., Jain, A.: Information fusion in biometrics. *Pattern Recogn. Lett.* **24**(13), 2115–2125 (2003)
44. Shakhnarovich, G., Moghaddam, B.: Face recognition in subspaces, *handbook of face recognition*, eds. stan z. li and anil k jain (2004)
45. Shruthi, B., Pooja, M., Ashwin, R.: Multimodal biometric authentication combining finger vein and finger print. *Int. J. Eng. Res. Dev.* **7**(10), 43–54 (2013)
46. Sim, T., Zhang, S., Janakiraman, R., Kumar, S.: Continuous verification using multimodal biometrics. *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(4), 687–700 (2007)
47. Simoens, K., Bringer, J., Chabanne, H., Seys, S.: A framework for analyzing template security and privacy in biometric authentication systems. *IEEE Trans. Inf. Forensics Secur.* **7**(2), 833–841 (2012)
48. Snelick, R., Uludag, U., Mink, A., Indovina, M., Jain, A.: Large-scale evaluation of multimodal biometric authentication using state-of-the-art systems. *IEEE Trans. Pattern Anal. Mach. Intell.* **27**(3), 450–455 (2005)
49. Soltane, M., Doghmane, N., Guersi, N.: Face and speech based multi-modal biometric authentication. *International Journal of Advanced Science and Technology* **21**(6), 41–56 (2010)
50. Sondhi, K., Bansal, Y.: Concept of unimodal and multimodal biometric system. *International Journal of Advanced Research in Computer Science and Software Engineering* **4**(6), 394–400 (2014)
51. Yao, Y.F., Jing, X.Y., Wong, H.S.: Face and palmprint feature level fusion for single sample biometrics recognition. *Neurocomputing* **70**(7–9), 1582–1586 (2007)
52. Zafeiriou, S., Tefas, A., Buciu, I., Pitas, I.: Exploiting discriminant information in nonnegative matrix factorization with application to frontal face verification. *IEEE Trans. Neural Netw.* **17**(3), 683–695 (2006)
53. Zhan, C., Li, W., Ogunbona, P.: Finding distinctive facial areas for face recognition. In: *International Conference on Control Automation Robotics Vision*, pp. 1848–1853 (2010)

54. Zhang, G., Huang, X., Li, S.Z., Wang, Y., Wu, X.: Boosting local binary pattern (Lbp)-based face recognition. In: *Advances in Biometric Person Authentication*, pp. 179–186. Springer (2004)
55. Zhang, Y., Sun, D., Qiu, Z.: Hand-based feature level fusion for single sample biometrics recognition. In: *2010 International Workshop on Emerging Techniques and Challenges for Hand-Based Biometrics (ETCHB)*, pp. 1–4. IEEE (2010)
56. Zhou, X., Bhanu, B.: Feature fusion of side face and gait for video-based human identification. *Pattern Recogn.* **41**(3), 778–795 (2008)
57. Zhou, X., Xu, G., Ma, J., Ruchkin, I.: Scalable platforms and advanced algorithms for IoT and cyber-enabled applications. *J. Parallel Distrib. Comput.* **118**, 1–4 (2018)
58. Zhou, X., Zomaya, A.Y., Li, W., Ruchkin, I.: Cybermatics: Advanced strategy and technology for cyber-enabled systems and applications. *Futur. Gener. Comput. Syst.* **79**, 350–353 (2018)
59. Zois, E.N., Tsourounis, D., Theodorakopoulos, I., Kesidis, A.L., Economou, G.: A comprehensive study of sparse representation techniques for offline signature verification. *IEEE Transactions on Biometrics, Behavior, and Identity Science* **1**(1), 68–81 (2019)

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Sidra Aleem¹ · Po Yang² · Saleha Masood¹ · Ping Li³ · Bin Sheng¹ 

Sidra Aleem
sidraaleem1234@gmail.com

Po Yang
P.Yang@ljmu.ac.uk

Saleha Masood
salehamasood08@gmail.com

Ping Li
lipingfire@ieee.org

¹ Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

² Department of Computer Science, Liverpool John Moores University, Liverpool, England

³ Department of Computing, The Hong Kong Polytechnic University, Kowloon, Hong Kong