1. Introduction

Biometrics is a pattern recognition technique recognizing humans based on their physiological or behavioral traits which has been gaining research interest from last few decades [1,2]. A typical biometric authentication system involves data acquisition, feature extraction followed by matching and decision making. The main steps carried out in biometric systems are shown in Fig. 1.

Unimodal biometric system is based on a single trait and it suffers from various limitations such as spoof attacks and several others as stated in literature [3]. While a multimodal biometric system is created by fusing various unimodal systems to ensure high performance of such biometric system as the evidences from different sources are combined together to avoid limitations of unimodal system as shown in Fig. 1. These sources can be from different sensors based on a single biometric or different entities based on a single biometric, like palm feature vector obtained from left and right hands or multiple biometric traits. The information retrieved from individual systems is combined using various schemes such as capturing same information using multiple sensors. For example, the audio samples of an individual are captured using a Hi-tech microphone and an Iphone. Also various traits of an individual are combined to yield a multimodal biometric system. The cost of these systems is high since each modality requires separate sensor and data acquisition phase. For example, face and fingerprint of an individual used for designing a multimodal system. Also, information from similar trait can be combined, for example, palm feature vectors obtained from left and right hands of same individual. These systems are cost-effective, since they need neither new sensors nor new algorithms for feature extraction. The single trait can be processed using multiple algorithms.

Multimodal biometric system shown in Fig. 2 improves the performance of the system by reducing system error rates and produces better recognition rate using some fusion techniques. It increases the robustness of the system in dealing with users who cannot be enrolled with a specific modality and also increases the resistance to spoof attack [4]. By having more than one biometric modality, the reliability of the system improves by making up the loss of one modality with another. Fusion techniques are used to aggregate the feature data extracted from the modalities to produce the improved recognition rates. In multimodal biometric system, complementary information that overcomes the drawbacks of unimodal biometric system is fused. Owing to multiple and independent data obtained from various modalities the multimodal biometric system is more reliable and has higher verification rate and improved accuracy. As a number of modalities are...
at our disposal, the choice of suitable biometric trait to authenticate a person becomes an easy task, thus addressing the issue of non-universality. If the biometric sample acquired is not of sufficiently good quality, then the samples from other sources can be banked upon to provide ample discriminatory information to ensure reliable decision-making. Noise in the sensed data from multiple traits has a lesser probability of affecting the performance of a biometric system. Multimodal biometric systems facilitate the choice of the modalities in a given situation. They add more flexibility to the enrollment procedure during user authentication.

The main issues and challenges in the design of multimodal biometric are: non-availability of multimodal biometric data; choice of modalities and choice of fusion technology. Apart from these issues, the performance evaluation of biometric systems is essential in high security applications like defense, government sector, airports, and forensics and also in commercial applications like access control, mobile computing. Biometrics based authentication involves personal data; it may be possible that data collected may be used for some unanticipated purpose. Privacy concerns are related to data collection, unauthorized use of recorded information and improper access to biometric records.

Based on the issues of unimodal biometrics, several multimodal biometric methods are developed and tried with many fusion mechanisms such as face and fingerprint using match score fusion [5]; face and speech [6]; face, voice and lip movement [7]; face, fingerprint and hand geometry at the matching score level [8]; palm-print, hand shape and knuckle print using feature level fusion [9]; integrating iris, face and palmprint [10]; finger vein, fingerprint and the finger-knuckle-print using Kernel Fisher analysis [11]; and fusion of input gestures from multiple modalities was presented [12]. In [13], a biohashing based fusion of palmprint and palm vein is suggested which shows the conversion of features in biocodes for security purposes. Another fusion approach of palmprint and palm vein is presented in [14] which presents opponent-processing and dual-tree complex wavelet transform (DTCWT) are more effective than discrete wavelet transform (DWT) and shift invariant discrete wavelet transform (SIDWT). In [15], biometric approach of car anti-theft system is suggested using fusion of palmprint and palm vein. The conventional sum rule, product rule, max rule, min rule, median rule and majority voting are presented by fusing evidences from multiple classifiers [16]. In 2011, a multimodal biometric system is presented by combining palm-print and finger knuckle-print at the matching-score level [17]. The score level fusion of knuckle, palmprint, and hand geometry is implemented using t-norms [18]. In 2015, hybrid fusion of score level and adaptive fuzzy decision level [19] is also presented which proves that single modality from all four finger knuckle-print can be combined in multimodal fashion.

In this paper, an attempt has been made to improve the performance of a biometric system by using palmprint, dorsal hand vein pattern and a novel biometric modality “palm-phalanges print”. A fusion of three different biometric modalities which includes palm-print (PP), palm-phalanges print (PPP) and dorsal hand vein pattern (DHV) has been proposed. We also present score level fusion of PP-PPP, PP-DHV, PPP-DHV and PP-PPP-DHV strategies.

This paper is organized as follows. Section 1 gives an introduction of unimodal and multimodal biometric systems. Section 2 explains the process of database collection. Section 3 gives the idea of “palm-phalanges print” and its feature extraction. Section 4 and Section 5 present the method of palmprint and dorsal hand vein feature extraction respectively. Sections 6 and 7 demonstrate the method of fusion. Section 8 presents the stage wise steps followed in proposed work with graphical representation. Simulation results are shown in Section 9. The paper is concluded in Section 10.

2. Database acquisition

First, we have collected a new anterior hand database of 50 individuals with 500 images (50*10 samples) in biometric laboratory of the institute referred to as NSIT Palmprint Database 1.0 by using a NSIT palmprint device shown in Fig. 3. Then, this database is used to extract region of interest (ROI) from anterior hand images. The image procurement setup designed here is inherently simple, fast and user-oriented so as to make the data acquisition process simple. Same database is used to extract ROI of palm-phalanges print as shown in Fig. 4. Here, the individuals do not have to undergo the distress of using two different sensors since the palmprint and palm-phalanges print features can be captured from the same image, using NSIT palmprint device, at the same time. In acquired database, fingers are not touching each other but all sample positions are kept varying so as to make the system position invariant and hence optimizing the practicality of the system.

Fig. 1. A typical biometric authentication system involves data acquisition, feature extraction followed by matching and decision making.

Fig. 2. A general multimodal system.
2.1. NSIT palmprint device

NSIT palmprint device consists of a charge-coupled device (CCD) array, two mirrors, a glass plate, a wooden box, a lamp, a stepper motor, a belt, a power supply and a control circuitry. A wooden box with a glass plate at its top is used to cover all the required circuitry of the device shown in Fig. 3. Data acquisition starts when you place your palm on a glass plate, and a CCD camera takes a picture. The palm is placed on the glass plate which provides a uniform background. A lamp is used to reflect light at the palmprint. CCD is used to convert photons (light) into electrons (electrical charge) to form an image of the subject such as palmprint. An analog-to-digital converter is used to process the analog electrical signal to generate a digital representation of this image. The image of the palm reaches the CCD array through a series of mirrors, filters and lenses. The entire mechanism (two mirrors and CCD array) make up the image capture panel (ICP). The ICP is moved slowly across the document by a belt that is attached to a stepper motor. The image of the palmprint is reflected by an angled first mirror to second mirror then on the CCD array. The resolution of these images is less than 200 dpi with 600*400 pixels and all these images are available in bitmap format.

2.2. Nomenclature

After assimilating the anterior hand samples, each sample is tagged with a unique identification code.
(Person ID) : This represents sample of the individual which is unique to every person.
(Sample number) : This represents sample number varied from 01 to 10.

3. Palm-phalanges print

Human hand with its two dorsal and anterior sides consist of phalanges joint, due to its unique pattern can be claimed to be used in personal identification. In its physical anatomy, four fingers, i.e. little, ring, middle and index, slightly differ from thumb. In palmer side of fingers, the skin folding is known as phalanges. Palmprint is the middle region of the ventral part of the hand. The skin in this area contains ridges like structure (dermal papillae) to increase friction which are also present on the fingers, i.e. fingerprints. The pattern in the skin folding of phalanges joint contains broad area with parallel lines showing the unique information about the individual. This joint can be used as biometric modality and termed as palm-phalanges print.

3.1. Aging effects and stability

When it comes to security and authentication systems, the biometric aging effect and stability comes under consideration when user uses the system after a time span of some years. Although most of the biometric traits are physiological or behavioral, they remain stable for small time period say 5–8 years. But if this time span is longer, i.e. more than a decade then we need to perform various tests and experiments in order to validate our biometric system. A survey of aging effects on biometric systems is presented in [20]. To study the age effect on palm-phalanges, a database of long term elapse is required. But we do not have this type of database. Therefore to study this effect, we assume that skin of phalanges and palm surface is similar and hence presents the stability of phalanges in term of palmprint. The anterior surface of palm and fingers is much thicker than dorsal surface. One of the upshots of the thinner dorsal hand skin in elderly adults is that it is more fragile, drier, and heals more slowly after injury. One of the more obvious signs of dorsal skin aging of the hand is wrinkling and loss of elasticity. But, aging changes in the thick skin of the anterior surface of the palms and fingers are less apparent than those in the thin skin of the dorsal hand [21]. Most of the anterior surface remains same and heal up earlier than that of dorsal skin. There is less wrinkling and loss of elasticity in palmer side and hence the recognition system based on this modality is less susceptible to aging effect and is highly stable. It was also concluded in [22] that stable recognition system can be sustained over aging effect if multiple fusion is to take place.

3.2. Palm-phalanges print feature extraction

Novel palm-phalanges print is presented here which, in turn, can be used as biometric modality as shown in Fig. 5. For feature extraction, postures of all fingers must be same so as to capture the same set of information from joints. Also in database, variations in sample position are made at the time of data acquisition. To make the hand samples rotation invariant, some pre-processing is required. For this, the coordinates of five fingertips and finger valleys considered as the key points and the centroid from each hand image is extracted. With the help of centroid and fingertips, the
image is rotated such that the line joining the tips of the index finger and ring finger becomes horizontal. Then coordinates of finger valleys are used to crop the region of interest (ROI) of each finger of size 200*100 as shown in Fig. 6. For enhancement of pattern, adaptive histogram equalization [23] based on Rayleigh distribution is employed on all the ROI.

After pre-processing, the enhanced ROIs are partitioned into non overlapping windows of size 20*10 each. Thus a total of 100 windows are created from each ROI. Then the Gaussian membership function based feature of palm-phalanges (GMF) of length 100 is extracted from each window for feature extraction.

Then the Gaussian membership function is used to extract the features from each window because of its robustness and property of providing system with less degree of freedom [24]. The feature vector so obtained has a length of 100. The formula for GMF [25,26] based feature extraction to extract GMF feature \( a_i \) from ith window are expressed in Eqs. (1 and 2).

\[
\begin{align*}
    u_i &= \frac{\exp\left(-\frac{(x_k - \bar{x})^2}{2\sigma^2}\right)}{2\pi} \\
    a_i &= \frac{1}{K} \sum_{k=0}^{K} x_k u_i
\end{align*}
\]

where \( x_k \) is the pixel value at kth point of the window, \( \bar{x} \) is mean pixel value and \( \sigma \) is the standard deviation of the window, \( u_i \) is the membership function and \( a_i \) is the feature obtained from the \( i_{th} \) window.

Along with GMF, the features based on statistical mean and average absolute deviation (AAD) which have been reported in literature for feature extraction are also extracted. Average absolute deviation algorithm is defined as:

\[
a_{AA} = \frac{1}{K} \sum_{k=0}^{K} |m(x, y) - m|
\]

where \( K \) is the number of pixels in the image, \( m \) is the mean of the image and \( m(x, y) \) is the value at point \( (x, y) \). At last, all four finger phalanges are fused to form a vector of size 1 x 400.

4. Palmprint feature extraction

For feature extraction of palmprint [27–29], similar procedure of extracting finger valleys are used to crop the region of interest (ROI) from an image of size 400*350. For enhancement of the palm pattern, adaptive histogram equalization based on Rayleigh distribution is employed on all the ROI.

After pre-processing, the enhanced ROI images of size 400*350 are partitioned into non overlapping windows of size 40*35 each as shown in Fig. 7. Thus a total of 100 windows are created from each image. Then GMF, mean and AAD features of length 100 are obtained using Eqs. (1, 2 and 3).

5. Dorsal hand vein feature extraction

Dorsal hand vein is a physiological biometric trait used for human identification [30–32]. Here Bosphorus Hand Vein Database [33] is used for dorsal hand vein feature extraction. Same procedure of extraction of ROI of palmprints as explained in Section 4 is followed here. After extraction of ROI of size 180*180, same enhancement and preprocessings are applied to non overlapping windows of size 18*18 each as shown in Fig 8. Then GMF, mean and AAD features of length 100 are obtained from each window.

6. Fusion of biometric modalities

A unimodal biometric system that uses only a single biometric modality in making personal authentication is often not able to meet the desired performance. Authentication based on multiple biometric modalities referred to as multimodal biometric authentication offers greater reliability and improved performance. It has been proved in the literature that the information from multiple biometric traits can achieve the stringent security requirements of the real-world applications. The various biometric modalities have their own advantages and disadvantages in terms of accuracy and user acceptance. The biometric application suggests the choice of a specific biometric identifier. A multimodal biometric system which uses multiple traits enables a biometric system to operate effectively in different applications and environments.

6.1. Fusion levels

Fusion of biometric modalities can be done at various levels. There are mainly four types of fusions in a multimodal biometric system:

1. Score level
2. Feature level
3. Sensor level
4. Decision level

7. Score level fusion

The score level fusion also called as confidence level fusion refers to combining the matching scores obtained from different classifiers. The block diagram depicting score level fusion is shown in Fig. 9. Each biometric modality provides a similarity score indicating the proximity of the test feature vector with the template feature vector. The fusion at score level is most appropriate
approach to multimodal biometrics and is most popular. The advantages of score level fusion are as follows: The matching scores (genuine and imposter) from the existing and proprietary unimodal systems can be easily utilized in a multimodal biometric system. The information (i.e. the match score) from prior unimodal evaluations of a biometric system can be used and this avoids live testing. The matching scores contain next level of rich information after the features of the input pattern. The scores generated by different matchers are easy to access and combine.

7.1. Normalization methods

Different modality scores must be transformed to common domain so that their combinations proved to be meaningful, this process of scaling is called normalization [34]. The score normalization technique used in this work is Min Max Normalization which is the simplest normalization and is best suited for the case where the bounds (maximum and minimum values) of the scores produced by a matcher are known. The minimum and maximum scores are shifted to 0 and 1, respectively. Let $s_k$ denotes a set of matching scores where, $k = 1, 2 \ldots n$ and $s_k$ denote normalized score. Then, the normalized score is given as:

$$s_{kn} = \frac{s_k - \text{min}}{\text{max} - \text{min}}$$

8. Stage wise steps followed in proposed work

Steps followed in proposed work are discussed in the form of algorithm as below. A general block diagram of the proposed work is shown in Fig. 10. The graphical representation of this work mainly includes three stages. First stage explains the procedure of data collection which consists of NSIT Palmprint Database 1.0 and Bosphorus Hand Vein Database. Stage 2 explains the pre-processing of both the databases to make hand rotation invariant. In stage 3, feature extraction, classification and fusion are explained.

Stage 1 Collection of database
1. Designing of NSIT palmprint device.
2. Anterior hand database containing palm-phantanges print (PPP) and palmprint (PP) is collected through NSIT palmprint device.
3. Dorsal hand vein (DHV) database is procured from standard source for fusion.

Stage 2 Pre-processing
1. Hand samples of both the databases have position difference. To make hand samples rotation invariant, coordinates of fingertips and finger valleys and centroid are calculated.
2. Hand samples are straightened using fingertips and centroid.
3. Next ROIs are extracted using finger valleys of straightened hand samples.

Stage 3 Feature extraction, classification and fusion
1. Apply adaptive histogram equalization (AHE) on each ROI, based on Rayleigh distribution.
2. Feature extraction methods are applied using GMF based features, AAD features and mean features.
3. Classification is performed using KNN, SVM and Random forest. Calculate scores and obtain ROC for individual modality.
4. Apply score level fusion using different rules with PP-PPP, PP-DHV, PPP-DHV and PP-PPP-DHV strategies. Performance analysis is done on the basis of GAR Vs FAR in receiver operating characteristics (ROC).

9. Simulation and results

Several popular and well known classification algorithms to evaluate the performance of biometric system are investigated here. \(K\)-nearest neighbor (K-NN) classifier with Euclidean distance is trained with features obtained from each biometric modality with k-fold cross-validation. The SVM [35] using polynomial kernel gave much better results than those from radial basis function. Therefore, only the results from polynomial kernel are reported. For degree-d polynomials, quadratic polynomial \(d=2\) is reported as it is giving better results than \(d=1\), and \(d=3\). Random forest (RF) [36], a pattern recognition method based on “ensemble learning” strategy, is also reported here with learning rate as 0.1.

9.1. Validation of palm-phantanges print as a biometric modality

To validate the use of palm-phantanges as a biometric modality, we combined the features obtained by four fingers and recognition results are evaluated using K-NN, SVM and RF. The features obtained are mean features, AAD and GMF as discussed in Section 3.2 and their recognition results are shown in Table 1. This table also shows the performance of corresponding classifier with respective feature.

The best performance of recognition for different modalities is achieved with K-NN classifier which is inherently simple and does not require training phase. While performance of SVM is also good when the second order polynomial kernel is used. However, the performance of random forest has been not so good. On the basis of feature extraction method, GMF method outperforms AAD and mean in most of the classifiers. Recognition results of palmprint (PP) and palm-phantanges print (PPP) for NSIT Palmprint Database 1.0 and Dorsal Hand Vein (DHV) for Bosphorus Hand Vein Database are also tabulated in Table 1.

First, it is shown that K-NN is worth considering and achieved good overall performance than SVM and RF. Second, as compared to SVM, both K-NN and RF are very similar and well understood. SVM however is, more appealing theoretically and in practice, its strength is its power to address non-linear classification task. However, most of the tasks examined here were not like that. We also observed that results highly depend on the adopted methodology.

Fig. 11 shows the receiver operating characteristics (ROC) from the respective features and their area under curve (AUC). The score
Fig. 10. A general block diagram of the proposed work.

Table 1
Comparative performance evaluation in percentage for palm-phalanges print, palmprint and dorsal hand vein pattern using different classifiers.

<table>
<thead>
<tr>
<th>Modality</th>
<th>KNN</th>
<th>SVM</th>
<th>Random forest</th>
<th>Feature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palm-phalanges print (PPP)</td>
<td>97.1</td>
<td>96.3</td>
<td>95.2</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>97.6</td>
<td>97.8</td>
<td>94.4</td>
<td>AAD</td>
</tr>
<tr>
<td></td>
<td><strong>98</strong></td>
<td><strong>98</strong></td>
<td><strong>96.4</strong></td>
<td><strong>GMF</strong></td>
</tr>
<tr>
<td>Palmprint (PP)</td>
<td>90.6</td>
<td>89.1</td>
<td>82.3</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>89.8</td>
<td>86.6</td>
<td>82.3</td>
<td>AAD</td>
</tr>
<tr>
<td></td>
<td><strong>94</strong></td>
<td><strong>91.1</strong></td>
<td><strong>82.8</strong></td>
<td><strong>GMF</strong></td>
</tr>
<tr>
<td>Dorsal hand vein pattern (DHV)</td>
<td>90.1</td>
<td>81</td>
<td>81</td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>80.3</td>
<td>84.1</td>
<td>81.1</td>
<td>AAD</td>
</tr>
<tr>
<td></td>
<td><strong>92</strong></td>
<td><strong>90.8</strong></td>
<td><strong>85.6</strong></td>
<td><strong>GMF</strong></td>
</tr>
</tbody>
</table>

Fig. 11. Receiver operating characteristic of mean features, average absolute deviation (AAD) and Gaussian membership features (GMF) of palm-phalanges print.

obtained are used to verify the performance of the recognition system using ROC curve between the Genuine acceptance rate (GAR) and false acceptance rate (FAR). It is seen in ROC that GMF features outperform other feature extraction methods. For GMF features, at 0.1 FAR, GAR is 94.25% and with FAR = 1, GAR is 96.2%. This also validates good performance of palm-phalanges print. For AAD and mean, GAR is low as compared to GMF.

For ensuring the clarity of plots and comparing other modalities, ROC of GMF based features of palmprint, palm-phalanges and dorsal hand vein are shown in Fig 12. It is seen that, at 0.1 FAR, GAR is 85% for PP, while for PPP, it is 94.25%, for DHV, it is 64.8%. Similarly at FAR = 1, GAR is 91.85% for PP, while for PPP, it is 96.2%, for DHV, it is 80.5%. This shows better performance of palm-phalanges trait over dorsal hand vein and palmprint.
9.2. Results of score level fusion

To compare the multimodal fusion, three modalities are combined in four PP-PPP, PP-DHV, PPP-DHV and PP-PPP-DHV ways using score level fusion. As the phalanges and palmaprint database are created using the same samples, this shows advantage of using same database in multimodal fusion. Phalanges was also fused with vein database which shows the compatibility of our database with others. Various conventional, ordered weighted averaging (OWA) operators and t-norm based fusion are performed here. Firstly, conventional sum and product rule of fusion operators are applied [16]. Secondly, Yager’s OWA based fusion is applied with appropriate weights [37]. Finally, Hamacher T-norm, Frank T-norm, Schweizer–Sklar and Dubois–Prade [38,18] are performed. Identification results of score level fusion of PP-PPP, PP-DHV, PPP-DHV and PP-PPP-DHV strategies are tabulated in Table 2.

ROC curves for Score level fusion of PP-PPP, PP-DHV, PPP-DHV and PP-PPP-DHV strategies for all the mentioned rules are shown in Fig. 13, Fig. 14, Fig. 15, Fig. 16, Fig. 17, Fig. 18 and Fig. 19, respectively.

On comparing the fusion rules, ROC of Figs. 13 and 14 shows worst performance. However the performance of other rules is highly comparable. Most of them reach to GAR = 99.6 at FAR = 1. Therefore to differentiate the performance of the fusion methods, area under the curve of ROC is calculated. It is seen in Fig. 20 that Hamacher T-norm converge to 99.76% at FAR = 1, but its AUC is 982. Yager’s OWA and Dubois–Prade shows similar performances. It is clear that frank T-norm fusion rule has AUC = 0.999 which is highest and converges to 100% more rapidly hence this outperforms the other. Also the score level fusion of PP-PPP-DHV performs better than other three. This shows that multimodal fusion is better than unimodal fusion. Fig. 21 shows the score distribution of Yager’s OWA, frank T-norm and Dubois–Prade. To show the high accuracy
of fusion, overlapping between the genuine and imposter distribution must be very less. In Frank t-norm, there is least overlapping between genuine and imposter score distribution thus ensuring higher performance.

10. Conclusion

In this paper, a new anterior hand database is created which is referred as NSIT Palmprint Database 1.0 by using a NSIT palmprint device. Along with this, a novel biometric modality “palm-phanalges print” has been proposed. Result shows that this modality can be used in recognition systems effectively. To further improve the performance of unimodal biometric system, fusion of palm-phanalges print with palmprent and dorsal hand vein is presented. In multimodal biometric system, complementary information is fused to overcome the drawbacks of the unimodal biometric systems. The results shown prove that the performance of multimodal biometric system is significantly improved as compared to the unimodal biometric systems. The fusion of palmprent and palm-phanalges print proved to be very efficient and effective to the users as well as to the recognition system because data for both the modalities is captured using same image before analysis is done. The score level fusion is attempted with PP-PPP, PP-DHV, PPP-DHV and PP-PPP-DHV strategies. Various conventional operators, Yager’s ordered weighted averaging (OWA) operator and t-norm fusion operators are used for score level fusion. On comparing the fusion rules, the convergence rate of ROC of frank T-norm fusion is better than other fusion rules thus outperforming others. While Hamacher T-norm, Yager’s OWA and Dubois–Prade shows the similar performance. Also the score level fusion of PP-PPP-DHV performs better than other three. This shows that multimodal fusion is better than unimodal fusion.

References


