Implementation of a Multimodal Biometrics Recognition System with Combined Palm Print and Iris Features

Rabab M. Ramadan, Elaraby A. Elgallad

Abstract—With extensive application, the performance of unimodal biometrics systems has to face a diversity of problems such as signal and background noise, distortion, and environment differences. Therefore, multimodal biometric systems are proposed to solve the above stated problems. This paper introduces a bimodal biometric recognition system based on the extracted features of the human palm print and iris. Palm print biometric is fairly a new evolving technology that is used to identify people by their palm features. The iris is a strong competitor together with face and fingerprints for presence in multimodal recognition systems. In this research, we introduced an algorithm to the combination of the palm and iris-extracted features using a texture-based descriptor, the Scale Invariant Feature Transform (SIFT). Since the feature sets are non-homogeneous as features of different biometric modalities are used, these features will be concatenated to form a single feature vector. Particle swarm optimization (PSO) is used as a feature selection technique to reduce the dimensionality of the feature. The proposed algorithm will be applied to the Institute of Technology of Delhi (IITD) database and its performance will be compared with various iris recognition algorithms found in the literature.

Keywords—Iris recognition, particle swarm optimization, feature extraction, feature selection, palm print, scale invariant feature transform.

I. INTRODUCTION

The iris is a thin, circular construction in the eye, liable for monitoring the diameter and size of the pupil and thus the quantity of light reaching the retina contains unique patterns that can be different under near-infrared lamination. The distinctive pattern in the human iris is shaped by 10 months of age, and residues unchanged throughout one’s lifetime. Iris recognition systems have become significant in the last decades and used in several applications. This is because the changeability of the iris between different individuals is huge and the iris image is somewhat insensitive to the angle of illumination and is also easy to use due to the circumstance that the iris can be captured in a less invasive manner [1], [2].

In everyday life, the hand aids as the most natural tool for a human being to perceive and reconstruct surrounding environments. Therefore, its prevalence in the field of biometrics is not surprising. Palmprint biometrics is relatively a new emerging technology that is used to identify people by their palm features. In fact, palmprint data may be collected with low cost devices and minimal cooperation from subjects is required. Moreover, various palm print features may be extracted to recognize a person, for example principal lines, ridges, wrinkles, texture and minutiae points may be cited [3].

In this paper, we present a bimodal biometric system merging iris and palm print biometric modalities. SIFT descriptor were extracted from both iris and palm print images for personal recognition. The main advantage in our system is that a preprocessing task is not essential to extract features (except to the segmentation of palm print region) since the SIFT method is invariant to image transformations [3]. Using SIFT matching, the comparison scores between the training and testing images from the IITD database are calculated.

In [4]-[6], Dr. Eberhart proposed PSO which is a computational model established on the theory of cooperative performance and swarming in biological populations inspired by the behavior of fish schooling and also the behavior of bird flocking. In recent times, PSO has been applied as an effective optimizer in a lot of fields like wireless network optimization, artificial neural networks, linear constrained function optimization, and data clustering [7]. In this paper, PSO is addressed for reduction of the number of features by using it as a feature selection technique. The proposed algorithm is applied to the IIT Delhi database and its performance will be compared with iris recognition algorithms found in the literature.

The paper is organized as follows: In Section II, we discuss the previous work on iris recognition, palm recognition and fusion in both levels: feature level and decision level. Section III motivates and gives a theoretical grounding to stages of recognition systems; in Section IV, we illustrate the system evaluation and its results, and finally, Section V discusses the results and concludes the paper.

II. RELATED WORK

There are many algorithms for matching available; they all basically depend on the shape, color and texture. Exact recognition of individuals can be done by extracting the most discriminating information present in an iris pattern. Only the important features of the iris must be coded so that comparisons between templates can be made. The template that is created in the feature encoding procedure will need an equivalent matching metric, which gives a measure of match between two iris templates [8]. The first one is the pioneer patent dealing with the general idea of the iris recognition
process. It was established by the ophthalmologists Flom and Safir [9] and it expired back in 2005. The second one, developed by the professor John Daugman, was used in protecting the iris-code approach and expired back in 2011. The procedure of iris recognition begins with the iris segmentation, then the transformation of the data to a 2-D polar coordinate system through the Rubber Sheet process which is proposed by the Daugman. The feature extraction process can be divided into three variants: the zero-crossing variant, the phase-based variant [10] and the texture analysis methods variant [11]. Daugman [12] extracted texture phase-based information by using multi-scale quadrature wavelets to get the signature of the iris with 2048 binary components [13].

In [14], the researchers calculated a zero-crossing representation of one-dimensional wavelet transform at diverse resolution levels of a concentric circle on the image of the iris to describe, the texture of the iris. In [15], the iris texture is represented with a Laplacian pyramid by Wildes et al., which is constructed with four dissimilar resolution levels and to determine whether the input image and the model image are both from the same class, they used the normalized correlation. Tisse et al. [16] analyzed iris characteristics using the analytic image constructed by the original image and its Hilbert transform. Emergent frequency functions for feature extraction were in essence samples of the phase gradient fields of the analytic image’s dominant components [17], [18]. Similar to the matching scheme of Daugman, they sampled binary emergent frequency functions to form a feature vector and used Hamming distance for matching. Park et al. [19] used a directional filter bank to decompose an iris image into eight directional subband outputs and extracted the normalized directional energy as features. Iris matching was done by computing Euclidean distance between the input and the template feature vectors. Kumar et al. [20] utilized correlation filters to measure the consistency of iris images from the same eye. The correlation filter of each class was designed using the two-dimensional. In [21], Hong and Smith proposed the octave band directional filter banks which are capable of both directional decomposition and an octave band radial decomposition.

For effective storage and retrieval of an eye image of the iris, an effective compression algorithm would have to be established. Contourlet transform is one of the directional transforms that can professionally extract the directionality features with multi-resolution capability from images that have textures with smooth contours. The contourlet transform achieves better performance than the wavelet transform by capturing the image singularities. In [22], [23], the conventional method for contourlet transformation, Laplacian pyramidal decomposition of images is applied in the first stage [23]. The band pass output result at several levels of Laplacian pyramids are analyzed using Directional Filter Banks (DFB) for extracting the angular information [24] but, due to the redundancy nature of Laplacian pyramidal representation of images, the conventional contourlet transform will have redundancy as well. This redundancy of conventional contourlet representation bounds the usage of contourlet transform for image compression applications [24], [25].

SIFT is a feature extraction method suggested by Lowe [26] allowing to extract persistent local feature points from an image. It became common for pattern recognition process [27] and object detection. Charfi [28] suggested a new approach for personal verification merging hand shape and palm print features extracted using the SIFT. This transform was enhanced its high variation and effectiveness in several applications particularly in object recognition and video tracking. These experiments on the IITD hand database establish promising outcomes by fusing at the matching level score the hand shape and palm print modalities [28]. Chen and Moon [29] fused Symbolic Aggregate Approximation (SAX) features and SIFT descriptors to extract palm print features. To improve precision, [30] combined SIFT and competitive code cores for palm print verification. In these studies, the strength of local invariant features is exposed by the fusion of various representations.

A touch less palm print identification method constructed on SIFT descriptors and sparse representation method is proposed [28], to excerpt palm print features of left and right palms. The fusion scheme is implemented at rank level using Support Vector Machines (SVM) classifier and probability distribution to produce the final identity of an individual. Experiments assessed on the CASIA palm print database and a proposed touch less REST (REGim Sfax Tunisia) hand database, notify favorable performances which are competitive to other standing palm print identification methods. In [27], SIFT are concerned with the localization of key points on the contour of the hand in order to describe the shape of the hand. To enhance as much as possible the number of matched key points and reduce false matches, two levels of SIFT matching refinement of hand shape process and one level for palm print process were used. The experiments were assessed on 1150 hand images of the IITD hand database, acquired from 230 subjects. The gained results are promising and competitive to other biometric systems based on hand images. A bimodal hand identification system was proposed based on SIFT descriptors, extracted from hand shape and palm print modalities [3]. A local sparse representation method was implemented in order to represent images with high discrimination. Moreover, fusion was performed at feature and decision levels using a cascade fusion in order to generate the final identification rate of the bimodal system. The experiments were applied on two hand databases: the Indian Institute of Technology of Delhi (IITD) hand database and the Bosphorus hand database containing, respectively, 230 and 615 subjects. The results of the suggested method are competitive to other widespread bimodal hand biometric methods over the two hand databases. The correct identification rate reaches 99.57% which is competitive compared to current systems.

A fusion context has the capability to improve the identification rate of the biometric system. Jain et al., [31] fuse the evidence of three different fingerprint matchers to determine the resemblance between two particulars sets. The three particulars matchers considered in their system are based
on the Hough transform. Han extracted seven indicated line profiles from preprocessed palm prints and three fingers and used wavelets to calculate low frequency information [32]. In a multi-biometric system, fusion is carried out at the abstract or decision level when the decisions obtained by the separate biometric matchers are available [33]. Methods for decision-level fusion using “AND” and “OR” rules [34], majority voting [35], weighted majority voting [36], Bayesian decision fusion and behavior knowledge space [37].

III. METHODOLOGY

The block diagram of the proposed system is shown in Fig. 1. The segmented palm print (ROI) palm print of IITD database [38] is used in this paper. The left hand is used for 230 subjects, five images for each with size 150x150 pixels in grayscale, while for iris images; IIT Delhi Iris Database (Version 1.0) is used [39].

A. SIFT

Feature extraction is the procedure of extracting the iris features; this process was employed SIFT [3] to extract constant local feature points from the input image. This is attained by using local maxima and minima of a difference of Gaussian function to select key locations in scale space. Matching each pixel to its neighbors is used to conclude maxima and minima of this scale space function [44] as in (1)-(3):

\[ L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \]  

\[ G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\left(x^2+y^2\right)/2\sigma^2} \]

and the Gaussian difference scale space is defined as:

\[ D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \]

The key points, which are insensitive to noise, and invariant to affine transformations must be considered, when extreme points locations are identified. Using gradient magnitude \( m(x, y) \) and the scale, around the location of a key point, a neighborhood is taken to form the key point descriptor, then the direction can be calculated in that section as in (4), (5).

\[ m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \]

\[ \theta(x, y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \]

where \( \theta(x, y) \) is the orientation of the key point.

B. PSO Algorithm

When PSO is implemented to answer an optimization problem, a swarm of computational elements, called particles, is used to discover the solution space for an optimum solution [4]-[6]. Each computation element, or particle, denotes a candidate solution and is identified with obvious coordinates in the D-dimensional search space. The position of the \( i \)th particle is represented as \( X_i = (x_{i1}, x_{i2}, \ldots, x_{iD}) \). The velocity of a particle (the rate of the change in position between the recent position and the succeeding) is represented as \( V_i = (v_{i1}, v_{i2}, \ldots, v_{iD}) \). Then the calculation of the fitness function is performed for each particle in the swarm and is matched to the fitness of the finest preceding result for that similar particle and is compared to the fitness of the finest particle between all particles in the swarm. After determining the two best values, the particles are evolved by upgrading their velocities and positions according to the following equations:

\[ V_{i+1} = \omega * V_i + c_1 * rand_1 * (p_{best} - X_i) + c_2 * rand_2 * (g_{best} - X_i) \]

\[ X_{i+1} = X_i + V_{i+1} \]

where \( N \) is the size of the swarm; \( p_{best} \) is the particle best reached solution and \( i = (1, 2, \ldots, N) \) and the global best solution is \( g_{best} \) in the swarm. Cognitive and social parameters are \( c1 \) and \( c2 \) that are bounded between 0 and 2. \( rand1 \) and \( rand2 \) are two uniform distribution random numbers \( U(0,1) \).

\[ -V_{max} \leq V_{i+1} \leq V_{max} \]

(the maximum velocity is \( V_{max} \)). To control the stability of the search algorithm between exploration and exploitation, the...
Inertia weight \( \omega \) is used. To reach the termination condition (maximum number of iterations \( K \)), the recursive levels will continue.

**Binary PSO and Feature Selection**

In [6], [11], a binary PSO algorithm has been established. In the binary version, the particle’s location is coded in the formula of a binary string that imitates the chromosome in a genetic algorithm. The velocity of the particle is used as the probability distribution for the position equation. The updates equation becomes as in (7).

\[
\text{If } \text{rand}_{3<\text{rand} < 1} \text{ then } X_{i}^{t+1} = 1; \text{else } X_{i}^{t+1} = 0 \quad (7)
\]

A bit value of \{1\} in any dimension in the position vector indicates that this feature is selected as an essential feature for the next generation, whereas a bit value of \{0\} indicates that this feature is not selected [6].

A binary PSO algorithm will be deployed for feature selection. The task of the binary PSO algorithm is to seek the most demonstrative feature subset in the extracted features.

**C. Fusion in Feature Level**

In recent times, multimodal methodologies for personal recognition have developed in order to improve the efficiency of unimodal biometric systems using a single biometric feature [40]-[42]. Certainly, the fusion of two or more biometric modalities is offered as a promising policy to improve the precision of a biometric system. In the feature fusion level, several feature vectors are concatenated to compose a single and higher feature vector [42]. Feature level fusion involves fusing the evidence presented by two biometric feature sets of the same individual [43]. A simple scheme would be to take the average of the two feature vectors corresponding to the two instances of the biometric signal and use the average feature vector as the new template. In this paper, we introduced a new technique for feature fusion level using PSO to select the most representative feature subset through the extracted feature vectors.

**D. Cross Validation**

In the proposed system, SVM is used as a classifier. SVM is founded on cross-validation technique. In order to estimate predictive models by dividing the original sample into a training set to train the system, and a test set, to evaluate it.

SVM is one of Maximum Margin Classifiers and is based on the Structural Risk Minimization (SRM). SVM map the input vector to a higher dimensional space where a maximal separating hyperplane is constructed. Linear support vector machine is initially framed for binary classification. In order that the proposed system has 230 subjects which are observed on the Structural Risk Minimization (SRM). SVM map the input vector to a higher dimensional space where a maximal separating hyperplane is constructed. Linear support vector machine is initially framed for binary classification. In order that the proposed system has 230 subjects which are observed.

Assume that the training data set and its labels is \((x_n,y_n), n=1,...,N, x_n \in \mathbb{R}^D, t_n \in \{-1, +1\}\), SVMs learning consists of the following constrained optimization:

\[
\begin{align*}
\min_{w,\xi_n} & = \frac{1}{2} w^T w + C \sum_{n=1}^{N} \xi_n \\
\text{s. t.} & \quad w^T x_n t_n \geq 1 - \xi_n \quad \forall_n \\xi_n \geq 0 \quad \forall_n
\end{align*}
\]

where \( \xi_n \) are the slack variables, \( w \) is the vector of coefficients, and \( C \) is the capacity constant.

The unconstrained optimization problem as in (9):

\[
\sum_{n=1}^{N} \max(1 - w^T x_n t_n , 0) \text{min}_{w_r} = \frac{1}{2} w^T w + C \quad (9)
\]

L1-SVM is not differentiable, the L2-SVM is used which minimizes the squared hinge loss as in (10):

\[
\min_{w_r} = \frac{1}{2} w^T w + C \sum_{n=1}^{N} \max(1 - w^T x_n t_n, 0)^2 \quad (10)
\]

To predict the class label of a test data \( x \):

\[
\text{arg} g_t \text{max}(w^T x) \quad (11)
\]

To extend SVMs for multiclass problem, one-vs.-rest approach is used [44]. Representing the output of the \( k \)-th SVM as in (12):

\[
a_k(x) = w^T x \quad (12)
\]

**E. Score Fusion**

The voting process [44] will be influenced by the class membership probability array that was extracted from the SVM classifier. In order to identify human identity for each feature vector, the index of the maximum value of each column of this array is extracted:

\[
X_{k,i} = \text{arg}_i \max \left( Y_{k,i} \right) \quad (13)
\]

where \( X \) is the output array that contains the index of test images for all features fusion techniques used in the system, \( Y \) is the class membership probability array, \( n \) is the number of the test images, and \( k \) is the number of features extraction fusion techniques.

Then, the voting technique is run to find the most frequent values in the output array \( X \).

\[
Z = \text{mode}(X_{k,i}) \quad (14)
\]

where \( Z \) is the class number array of the test images.

**IV. EXPERIMENTAL RESULTS**

The fusion in the feature level proposed in this paper is based on using the binary PSO algorithm that has been developed in [4]. The binary PSO algorithm’s task is to look for the most representative feature subset through the extracted features by applying the proposed algorithm to the selected feature vectors extracted. The search heuristics in PSO is iteratively adjusted guided by a fitness function defined in terms of maximizing class separation.
The database used in this paper is the IIT Delhi Database. For palm print images, the IIT Touchless Palmprint Database version 1.0 is used, while for iris images [38], the IIT Delhi Iris Database (Version 1.0) [39] is used. The Touchless Palmprint image database consists of hand images collected from the students and staff at IIT. This database of images was saved in bitmap format, and contains both left and right hand images for 230 subjects. The subjects in the database are in the age group of 14 years to 56 years old and voluntarily contributed at least five hand image samples from each hand. Automatically segmented and normalized palm print regions are also made available, in addition to the acquired whole hand images. The segmented palm print ROI of IITD database is used in this paper. The left hands are used for 230 subjects, five images each with size 150x150 pixels in grayscale. The IIT Delhi Iris Database mainly consists of the iris images collected from the students and staff at IIT Delhi, New Delhi, India. This database has been acquired in Biometrics Research Laboratory during Jan-July 2007 using JIRIS, JPC1000, digital CMOS camera. The currently available database is from 224 users, in bitmap format. The age of all the subjects in the database is in the range of 14-55 years (176 males and 48 females). The database of 1120 images is arranged into 224 different folders, each has an integer identification. The resolution of these images is 320x240.

The proposed algorithm was found to generate excellent recognition results with less selected features (1010 features). The features number has been reduced by 50% from its original number in the feature level fusion without PSO (2048), as shown in Table I. While computing the difference between the images, the new feature vector, with the less selected features, will decrease the computation time and increase its performance. The proposed algorithm achieves 99.55% recognition rate. After applying fusion at the score level (voting process), the overall recognition rate is 99.78%.

| TABLE I: COMPARISON OF RECOGNITION RATES AND NUMBER OF FEATURES FOR VARIOUS FUSION LEVELS |
|-----------------|----------|---------|---------|----------|
| Descriptor      | DB       | Features| RR      | SVM      |
| SIFT            | Iris IITD| 1024    | 98.44%  |          |
| SIFT            | Palm IITD| 1024    | 97.32%  |          |
| fusion          |          | 2048    | 100.00% | 99.78%   |
| fusion ave.     |          | 1024    | 99.78%  |          |
| fusion PSO      |          | 1010    | 99.55%  |          |

V. CONCLUSIONS AND RECOMMENDATIONS

In this paper, SIFT is used to extract features for both iris and palm print images. For fusion in the feature level, the PSO-based feature selection algorithm is used. It searches within the feature space for the optimal feature subset. The classifier performance and the numbers of selected features were considered for performance valuation using the IIT Delhi Database. Experimental results show the power of the PSO-based feature selection algorithm in producing outstanding recognition accuracy with the minimal set of selected features. The performance of the proposed algorithm is compared with the performance of average feature vector and was found to yield comparable recognition results of 99.55% with less number of selected features. The use of fusion in the feature level is hereby recommended for all firms and industries where security and personal identification is favored. Using voting at the decision level improves the recognition rate to 99.78%. The experimental results in this study show the efficiency of the recommended system.

REFERENCES

[21] P. Hong and M. I. T. Smith, "An octave-band family of
nonredundant directional filter banks”, IEEE proceedings.


