Efficient multimodal fusion for person authentication combining three biometric modalities

Fenghua Wang, Qiumei Zheng, Wenjie Meng
College of Computer and Communication Engineering
China University of Petroleum
Qingdao, China
fenghuawang@upc.edu.cn

Abstract—**Fusion of multiple biometric modalities for person authentication can compensate for the limitations in performance of each individual biometric system. This paper proposed an efficient person authentication method based on multimodal biometric fusion at the matching score level. For score level fusion, MPM (minimax probability machine) based fusion rule is applied to combine multiple matching scores, respectively from iris, facer and palmprint, to generate a single scalar score which is used to make the final decision. Experimental results show that the performance of the proposed method can bring obvious improvement comparing to the previous multimodal methods.**

Keywords- multimodal; biometric; MPM; fusion

I. INTRODUCTION

The increasing demand on enhanced security has led to an unprecedented interest in automated personal authentication based on biometrics. Biometrics refers to the technologies that use physiological or behavioral characteristics to authenticate a person's identity [1]. Many biometric modalities are used for verification and identification purposes. However, despite significant research, unimodal biometric systems often suffer from limitations due to non-universal biometrics traits, susceptibility to biometric spoofing or insufficient accuracy caused by noisy data [2], and hence, may not be able to achieve the desired performance requirement in real-world applications. This accuracy problem has recently been addressed through multimodal biometric (multi-biometric) fusion, which combines the information that is output by individual biometric. Multimodal biometric fusion has been shown empirically to improve the accuracy of biometrics-based verification (one-to-one comparison) and identification (one-to-many comparison)[3-5].

In this paper, we proposed an efficient person authentication method based on multimodal biometric fusion. The proposed multimodal scheme integrates three biometric modalities: iris, face and palmprint. Three biometric verifiers are fused at the matching score level. When fusing, instead of non-parametric algorithms and traditional parametric machine learning algorithms, an improved fusion strategy based on MPM (minimax probability machine) is employed. The most attractive property of MPM is that it can explicitly provide a worst-case bound on the probability of misclassification of future data when the mean and covariance matrix of the data are known. Moreover, in the proposed fusion strategy, the fusion of different scores is viewed as a combination problem, in which the score vector is combined to generate a fused single scalar score which is then used to make the final decision by a predefined decision threshold. This approach can increase flexibility and meet demands under more circumstances by adjusting the decision threshold. The experimental results on our constructed multimodal database prove the superiority of the proposed system.

II. OVERVIEW OF THE PROPOSED SCHEME

Face recognition, iris recognition and palmprint recognition all involve image preprocessing, feature extraction, matching and decision making. Multimodal fusion for three modalities can be done at the feature extraction level, the matching score level, or the decision level. Although feature sets usually contain more information data than the matching scores, features from different modalities are usually incompatible. Fusion at the decision level is thought to lack flexibility (due to the limited information from each classifier, e.g. no information on confidence of decisions). Thus, fusion at the score level is the most popular and frequently used method because of its good performance, intuitiveness and simplicity.
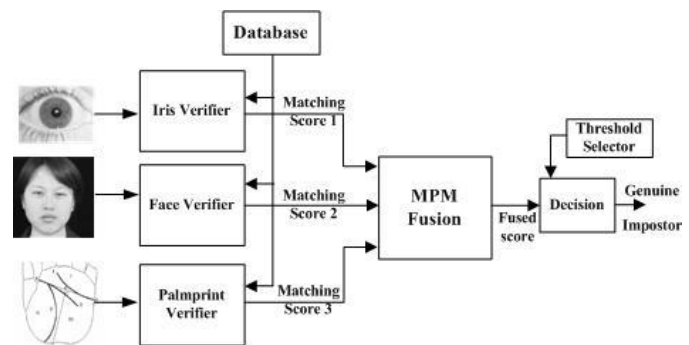

Figure 1.   Block diagram of the proposed scheme

Figure1 shows the block diagram of the proposed multimodal biometric authentication method integrating iris, face and palmprint. From Fig.1, it can be seen that the images of three modalities of a certain person waiting for being authenticated are first acquired and input corresponding verifiers to make a match with the stored

template in the database. Following the obtainment of matching scores, the matching scores from three modalities are normalized to transform into a common domain because the scores generated from different modalities are heterogeneous. At the fusion stage, the normalized scores are combined using the MPM-based fusion rule. At the decision step, a decision threshold is set to make a final decision. The decision threshold can be adjusted to meet demands of different application conditions.

III. RECOGNITION ALGORITHMS OF FACE, IRIS AND PALMPRINT

A. Face Recognition

Face recognition is an active area of research and numerous algorithms have been proposed for face recognition within the last several years. Among various algorithms, appearance-based approaches are the most popular. In our multimodal biometric system, the Laplacianface algorithm is employed in the face verifier part, which is a novel appearance-based face recognition algorithm[6].

In the Laplacianface algorithm, face images are mapped into a face subspace for analysis by using Locality Preserving Projections (LPP). LPP finds an embedding that preserves local information, and obtains a face subspace that best detects the essential face manifold structure. Laplacianface method can achieve better performance than other appearance-based methods such as Eigenface and Fisherface .

In our multimodal biometric system, the process of face recognition consists of the following stages:

1) Face image preprocessing. In this stage, the face images detected are normalized in order to reach scale and shift invariability. Face normalization is based on the position of two eyes and the distance between them. After normalization, eyes position and distance between two eyes are same. Then histogram equalization is applied to normalize the brightness level of face.

2) Training. In this stage, a set of training face images are collected and laplacianfaces are computed from the training set. The detailed process is as follows: First, the normalized face images are projected into the PCA subspace by throwing away the components corresponding to zero eigenvalue. Then Locality Preserving Projections is applied to reduce the number of features (dimensions). At last, the projection matrix can be represented as $W = W_{PCA}W_{LPP}$, in which each column of the projection matrix can be called as a laplacianface when it is transformed into two dimensions.

3) Recognition. The feature vector from an unknown facial image can be obtained by projecting the image into a face-space. In this process the image is represented as a linear combination of laplacianfaces and the feature vector is made of weightings associated with each laplacianface. The feature vector dimensionality was selected based on the classification experiments on the training set of the database.

4) Matching. The matching score between two face-feature vectors is calculated using the Euclidean distance in the matching phase. The formula can be denoted as:

$$d_E(v,u) = \sqrt{\sum_{i=1}^{k}(v_i - u_i)^2} \qquad (1)$$

In which, v and u are feature vectors of matching faces. k is the dimensionality of feature vector.

Following the above stage, the matching score of face verifier is obtained as the Euclidean distance.

B. Iris Recognition

The human iris is an annular region between pupil and sclera. Due to its high reliability and non-invasiveness, iris recognition is receiving increased attention. Among various algorithms, phase information based algorithm proposed by Daugman[7] is considered a very effective one, which used Gabor filters to extract phase structure information of iris. Our recent work show that better performance can be achieved by using 2D Log-Gabor filters to extract phase information[8]. So in the proposed multimodal scheme, our improved phase information algorithm using multi-scale 2D Log-Gabor is applied to generate the matching score of iris verifier. The detailed process is as follows:

1) Iris image preprocessing. Prior to feature extraction, the iris image needs to be preprocessed to eliminate uninterested information. The main preprocessing steps consist of localization of the inner and outer iris boundaries, localization of eyelid boundaries, transformation from polar coordinates to a fixed size rectangular image, mask generation and image enhancement.

2) Feature extraction and encoding. 2D Log-Gabor filters are employed to extract the phase information of iris. The iris image is divided into some blocks and the phase of each block can be extracted by using multi-scale 2D Log-Gabor filters[8]. The feature of iris can be described as a certain binary codes.

3) Matching. The difference between two iris was measured by their Hamming distance:

$$d_H = \frac{\sum [(codeA \otimes codeB) \bigcap (maskA \bigcap maskB)]}{\sum (maskA \bigcap maskB)} \qquad (2)$$

Where $\otimes$ denotes the Boolean Exclusive-OR operator (XOR), *maskA* and *maskB* denote two iris matching masks, respectively, "0" for the non-iris regions, and "1" for the iris regions; $\bigcap$ denotes the AND operator. Finally, the matching score of iris verifier is obtained as the Hamming distance.

C. Palmprint Recognition

The palmprint is a relatively new biometric feature used for automated personal authentication. In this work, a recognition algorithm (Laplacianpalmprint) similar to face recognition is applied. The process of computing the matching score between applicant and stored template is as follows:

1) Preprocessing. After the palmprints are captured by the CCD-based device, they should be aligned using the preprocessing technique. After preprocessing, the central part of the palmprint is cropped to represent the whole one.

2) Training and recognition. Similar to Laplacianface algorithm, a set of palmprints are as training data to train Laplacianpalmprint. When recognition, the palmprints being verified are projected into the trained palmprint subspace, and the feature vector can be obtained as the weightings associated with each Laplacianface.

3) Matching. The matching score between two palmprint feature vectors also can be calculated using the Euclidean distance.

IV. MATCHING SCORES NORMALIZATION

The matching scores generated from face, iris and palmprint verifier are heterogeneous because they are not on the same numerical range, which may negatively affect fusion performance. So normalization is required to transform these scores into a common domain before fusion at the matching score level.

A double sigmoid function is used for score normalization in this work. Given a set of matching scores d , the normalized score x is given by

$$x = \begin{cases} \dfrac{1}{1+\exp(-2((d-t)/t_1))} & d < t \\ \dfrac{1}{1+\exp(-2((d-t)/t_2))} & \text{otherwise} \end{cases} \quad (3)$$

Where t is the reference operating point and t_1 and t_2 denote the left and right edges of the region (i.e. the interval $(t-t_1, t-t_2)$) in which the function is near-linear. By using (3), the scores can be mapped to the [0,1] range.

V. MULTIMODAL FUSION AND DECISION

Following scores normalization, a multimodal score vector can be constructed. The next step is fusion at the matching score level. In this work, a flexible approach is adopted. The fusion of scores is viewed as a combination problem and the score vector is combined to generate a single scalar score which is used to make the final decision. This approach can meet demands under more circumstances by adjusting the decision threshold.

As to fusion strategies, rather than some conventional non-parametric learning fusion strategies such as sum, product, Fisher and SVM, MPM (minimax probability machine)is considered in our multimodal system[9]. The conventional fusion strategies have some disadvantages, for example, sum and fisher rule are often less effective in case of nonlinear separation problem. Although SVM can solve nonlinear separation problem, its performance is directly affected by the kernel function and its complexity is highly dependent on the size of training data, which make SVM not provide robust performance.

The most attractive property of MPM is that it can explicitly provide a worst-case bound on the probability of misclassification of future data when the mean and covariance matrix of the data are known.

Let the matching scores, provided by three modals, be combined into a multimodal score vector $\mathrm{d} = [x_1, x_2, x_3]^T$ ($x_1, x_2, x_3 \in R$). The design of a trained fusion scheme consists in the estimation of a function $f : R^2 \to R$ based on empirical data so as to effectively separate the fused scores $f(\mathrm{d})$ of genuines and impostors.

Suppose that two random vectors x and y represent two classes of data points with means and covariance matrices as $\{\bar{\mathrm{x}}, \Sigma_x\}$ and $\{\bar{\mathrm{y}}, \Sigma_y\}$, respectively, where x, y, $\bar{\mathrm{x}}$, $\bar{\mathrm{y}}$ $\in R^2$, and $\sum_x$, $\Sigma_y \in R^{2\times 2}$. Let x and y denote genuines and impostor, respectively.

The training sets, which consist of genuines and impostors, are already given. With the reliable estimations of $\{\bar{\mathrm{x}}, \Sigma_x\}$ and $\{\bar{\mathrm{y}}, \Sigma_y\}$ for two classes of data obtained from the training data, MPM attempts to determine an optimal hyperplane

$$\mathrm{a}^T \mathrm{z} = b \quad (\mathrm{a,z} \in R^2, \mathrm{a} \neq 0, b \in R), \quad (4)$$

which separates the data into genuines and impostors by minimizing the worst-case probability of misclassification.

By introducing the Lagrangian Multiplier, the optimization problem then becomes:

$$\max_{k,\mathrm{a}} k \text{ subject to: } \frac{1}{k} \geq \sqrt{\mathrm{a}^T \Sigma_x a} + \sqrt{\mathrm{a}^T \Sigma_y a}$$
$$\mathrm{a}^T (\bar{\mathrm{x}} - \bar{\mathrm{y}}) = 1 \quad (5)$$

Which allows us to eliminate k , and the problem can be transformed into:

$$\min_{\mathrm{a}} \left\| \Sigma_x^{1/2} \mathrm{a} \right\|_2 + \left\| \Sigma_y^{1/2} \mathrm{a} \right\|_2 \text{ subject to } \mathrm{a}^T (\bar{\mathrm{x}} - \bar{\mathrm{y}}) = 1 \quad (6)$$

This is a second order cone programming problem which can be solved using interior-point methods. Then we can obtain a_* and b_* , where a_* and b_* are the optimal values of a and b , and the optimal hyperplane also can be determined.

Given a new iris score vector z_{new} , a fused score s_T can be generated using the optimal hyperplane. s_T is defined as follow:

$$s_T = f(\mathrm{z}_{new}) = \mathrm{a}_*^T \Phi(\mathrm{z}_{new}) - b_* \quad (7)$$

Where Φ is called the kernel function, and Gaussian kernel function is adopted in this work.

Following the obtainment of the fused score s_T , the final decision can be made by the predefined threshold. In practical implementations, the decision threshold can be adjusted to reach different working points.

VI. EXPERIMENTS AND RESULTS

To evaluate the effectiveness of our proposed multimodal authentication scheme, a database containing iris, face and

palmprint samples is required. In this work, we construct a multimodal biometric database for our experiments based on UBIRIS iris database [10], ORL face database[11] and PolyU palmprint database[12].

The constructed multimodal database consists of 280 records corresponding to 40 subjects (7 records each subject) , and each record contains an iris image, a face image and a palmprint image. In our experiments, the 40 subjects are divided into two sets: 8 subjects (56 records) as training data to estimate the parameters of MPM, the remaining 32 subjects (224 records) as the test data to evaluate the performance of the trained system.

In the proposed multimodal scheme, the fusion of scores is viewed as a combination problem. And instead of some traditional fusion strategies such as sum, product, fisher and SVM applied in previous studies, the MPM-based score level fusion strategy is employed to generate a fused score for the final decision. In the experiments, we compared the proposed MPM-based fusion strategy with the traditional fusion strategies. The detailed comparison results are as follows.
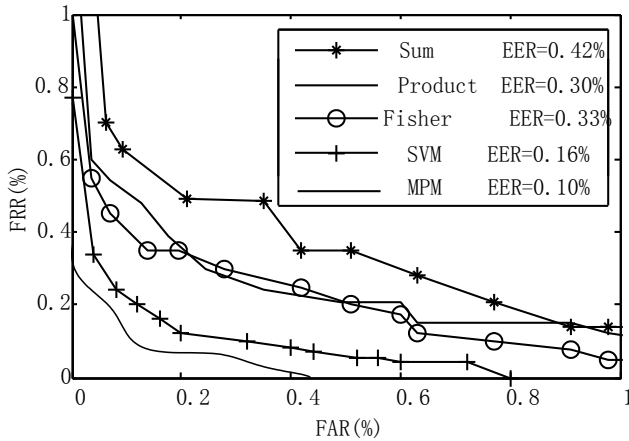

Figure 2.   ROC curves of different fusion strategies

In case that three modalities are all available, Figure2 gives the ROC curves for the multimodal biometric methods with different fusion strategies: sum, product, fisher, SVM and MPM. These strategies are employed at the matching score level to generate a fused score for decision. From the figure, we can see that: although these fusion strategies all can achieve good performance improvement compared with unimodal method, MPM based score level fusion rule can get the best accuracy and the most improvement among all fusion strategies, which prove the superiority of combination approach based on parametric learning fusion strategies.

VII.   CONCLUSIONS

This paper proposed an efficient person authentication method based on multimodal biometric fusion at the matching score level. For score level fusion, MPM (minimax probability machine) based fusion rule is applied to combine multiple matching scores, respectively from iris, facer and palmprint, to generate a single scalar score which is used to make the final decision. Experimental results show that the performance of the proposed method can bring obvious improvement comparing to the traditional fusion strategies.

REFERENCES

[1] A.K. Jain , A. Ross, S. Prabhakar. An introduction to biometric recognition, *IEEE Transactions on Circuits and System for Video Technology, Vol.* 14, 2004, 4-20.

[2] A.K. Jain , A. Ross. Multibiometric systems, *Communications of the ACM, Vol.* 47, 2004, 34-40.

[3] R. Raghavendra, BernadetteDorizzi, AshokRao. Designing efficient fusion schemes for multimodal biometric systems using face and palmprint[J]. Pattern Recognition, 2011,44(5):1076-1088.

[4] Norman Poh, Josef Kittler. A Unified Framework for Biometric Expert Fusion Incorporating Quality Measures[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2012, 34(1):3-18.

[5] Hofmann M., Schmidt S.M., et al. Combined face and gait recognition using alpha matte preprocessing [C]. ICB2012, 2012:390-395.

[6] H. Xiaofei , Y. Shuicheng , et al. Face recognition using laplacianfaces, *IEEE Transaction on Pattern Analysis and Machine Intelligence,Vol.* 27,2005,328-340

[7] J. Daugman. How Iris Recognition Works, IEEE Transactions on Circuits and System for Video Technology, Vol. 14, 2004, 21-30.

[8] F. Wang, J. Han. Iris recognition method using Log-Gabor filtering and feature fusion, *Journal of Xi'an Jiaotong University, Vol.*41, 2007, 889-893.

[9] Y. Liming, W. Laisheng, S. Yuhua etc. Simultaneous feature selection and classification via Minimax Probability Machine, International Journal of Computational Intelligence Systems, Vol. 3, 2010,754-760

[10] H. Proenca, A. Alexandre. *UBIRIS Iris Image Database*: http://iris.di.ubi.pt.

[11] A&AT, *The ORL Database of Faces*: http://www.cam-orl.co.uk/

[12] PolyU Palmprint Database: http://www.comp.polyu.edu.hk