ISSN 1392 – 124X INFORMATION TECHNOLOGY AND CONTROL, 2008, Vol. 37, No. 4

ROBUST MULTIMODAL BIOMETRIC AUTHENTICATION INTEGRATING IRIS, FACE AND PALMPRINT

Fenghua Wang, Jiuqiang Han

School of Electronics and Information Engineering, Xi'an Jiaotong University
710049 Xi'an, China

Abstract. Fusion of multiple biometric modalities for human authentication performance improvement has received considerable attention. This paper presents a robust multimodal biometric authentication scheme integrating iris, face and palmprint based on score level fusion. In order to overcome the limitation of the possible missing modalites, the multiple parallel support vector machines (SVMs) fusion strategy is applied, in which all possible modality combination cases are considered and each case has a corresponding SVM to combine the scores to generate a fused score for the final decision. Experimental results show that the proposed multimodal scheme is more robust and flexible, especially when some of the biometric modalities are unavailable.

Keywords: biometrics, human authentication, score level fusion, support vector machine.

1. Introduction

Biometrics refers to the technologies that use physiological or behavioral characteristics to authenticate a person's identity[1]. In recent years, the increasing demand on enhanced security has led to an unprecedented interest in automated personal authentication based on biometrics.

Biometric systems based on a single source of information are called unimodal systems. Although some unimodal systems have got considerable improvement in reliability and accuracy, they often suffer from enrollment problems 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. One way to overcome these problems is the use of multimodal biometric authentication systems, which combine information from multiple modalities to arrive at a decision. Some studies have demonstrated that multimodal biometric systems can achieve better performance comparing to the unimodal systems [2-7].

Although existing multimodal fusion techniques have been shown effectively to improve the accuracy of biometrics-based verification, they also face some limitations. For example, most existing multimodal fusion schemes, especially some single parametric machine learning fusion strategies, are based on the assumptions that each biometric modality is available and complete [4-6], so each registered person must be entered into every modality. Once a modality is unavailable or missed, the multimodal systems break down or the accuracy degrades. This may not be plausible and is very restrictive. Additionally, in existing multimodal fusion techniques, when the parametric learning fusion strategies are adopted at the matching score level [4, 5, 7], the fusion is viewed as a classification problem, in which the score vector is classified into one of two classes: "Accept" (genuine user) or "Reject" (impostor). However, the above approach seems to lack flexibility when different performance demands are required in real applications. Concerning these problems, some solutions are given in this work.

In this paper, we proposed a robust multimodal authentication scheme which addresses the problem mentioned above. 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 single machine learning fusion strategy, an improved fusion strategy based on a group of parallel support vector machines (SVMs) is employed. The parallel multiple SVMs cover all possible subsets of the biometric modalities being considered. The selector in fusion module can select an appropriate SVM for fusion from multiple SVMs according to the current available modalities, which can eliminate the limitation brought by the missing modalities. 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.

2. Framework of the proposed multimodal system

Face verifier, iris verifier and palmprint verifier 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.

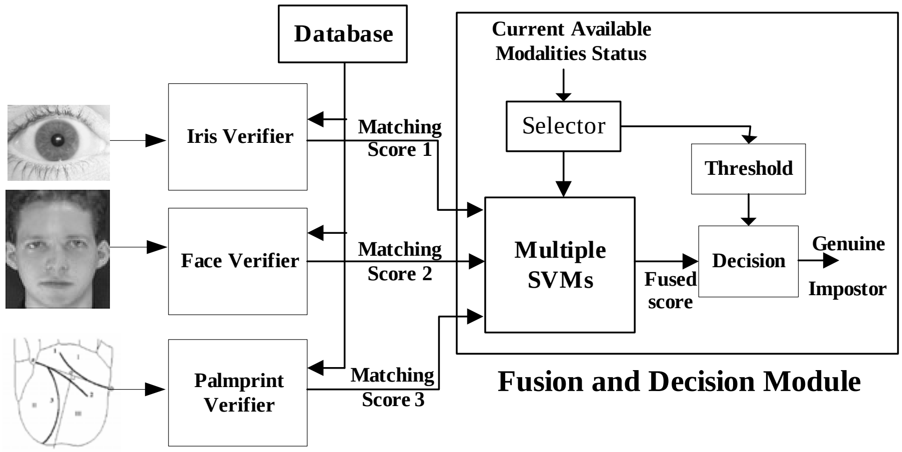


Figure1. Block diagram of the proposed scheme

Figure 1 shows the block diagram of the proposed multimodal biometric authentication method integrating iris, face and palmprint. From Figure 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 selector will select an appropriate SVM for fusion from multiple SVMs and corresponding decision threshold according to the current available modalities status. The selected SVM will combine multiple matching scores to generate a fused score for the final decision. At the decision step, the selected threshold is utilized to make a decision of genuine or impostor.

3. Biometric verifiers

3.1. Iris verifier

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[9] is considered a very effective one, which used Gabor filters to extract phase structure information of iris. Our recent work shows that better performance can be achieved by using 2D Log-Gabor filters to extract phase information [10]. 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:

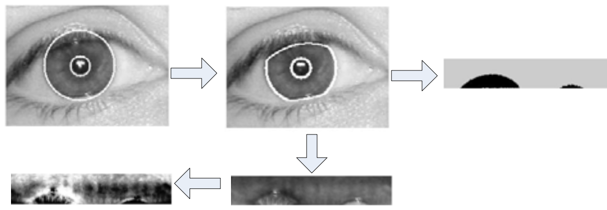


Figure 2. Steps involved in iris preprocessing

1) Iris image preprocessing. Prior to feature extraction, the iris image needs to be preprocessed to eliminate uninterested information. The main preprocessing steps, as illustrated in Figure 2, 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[10]. The feature of iris can be described as certain binary codes.
3) Matching. The difference between two iris was measured by their Hamming distance:

$$d_H = \frac{\sum[(codeA \otimes codeB) \cap (maskA \cap maskB)]}{\sum(maskA \cap maskB)}, \qquad (1)$$

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.

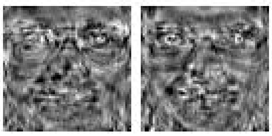
3.2. Face verifier

Face is an important biometric modality. Among various face recognition 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 newest appearance-based face recognition algorithm[8].

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. The process of face verifier consists of the following stages:

1) Face image preprocessing. The detected face images are normalized in order to reach scale and shift invariability. The histogram equalization is applied to normalize the brightness level of face. Figure 3(b) shows the preprocessed images.
2) Training. In this stage, a set of training faces 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 transformed into two dimensions. The examples of Laplacianfaces are shown in Figure 3(c).

(a) Original faces (b) Preprocessed faces



(c) Laplacianfaces

Figure 3. Images involved in face recognition

3) Recognition. The feature vector from a 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.
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:

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

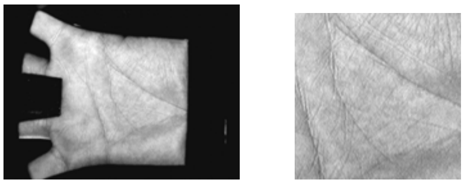
where v and u are feature vectors of matching faces. k is the dimensionality of feature vector.

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

3.3. Palmprint verifier

The palmprint is a relatively new biometric feature used for automated personal authentication[1]. 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 preprocessed to separate the fingers. After preprocessing, the central part of the palmprint is cropped to represent the whole one. Figure 4 shows a captured palmprint and its cropped images.



(a) Original palmprint (b) Cropped image

Figure 4. Images involved in palmprint recognition

2) Training and recognition. Similar to Laplacianface algorithm, a set of palmprints are as training data to train Laplacianpalmprint. When performing recognition, the palmprints being verified are projected onto the trained palmprint subspace, and the feature vector can be obtained as the weightings associated with each Laplacianpalmprint.
3) Matching. The matching score between two palmprint feature vectors also can be calculated using the Euclidean distance.

4. Fusion and decision

4.1. Score normalization

The matching scores generated from different verifiers are heterogeneous because they are not on the same numerical range, which may negatively affect fusion results. So the first step of fusion is score normalization to transform scores into a common domain.

A double sigmoid function is used for score normalization in this work. Given matching scores d of a certain verifier, 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.

4.2. Multiple parallel SVMs fusion strategy

After score normalization, a mutimodal score vector can be constructed. The next step is fusion at the matching score level. In previous researches, this step is often viewed as a classification problem, and by using some learning machines the score vector is directly classified into one of two classes: "Accept" (genuine user) or "Reject" (impostor). In this work, a more 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, for example, we can increase decision threshold to meet the circumstance claiming strictly low false rejection rate (FRR) and relaxed false acceptance rate (FAR), or we also can decrease threshold to meet the circumstance claiming strictly low false acceptance rate (FAR) and relaxed false rejection rate (FRR).

As to fusion strategies, rather than some conventional non-parametric learning fusion strategies such as sum, product, and Fisher, supervised parametric learning fusion strategies based on SVM are considered in our multimodal system. However, instead of the single static SVM fusion strategy, multiple parallel SVMs fusion strategy is proposed and utilized to overcome the limitation of previous techniques when some modalities are not currently available. The structure of the multiple SVMs fusion strategy is described in Figure 5.

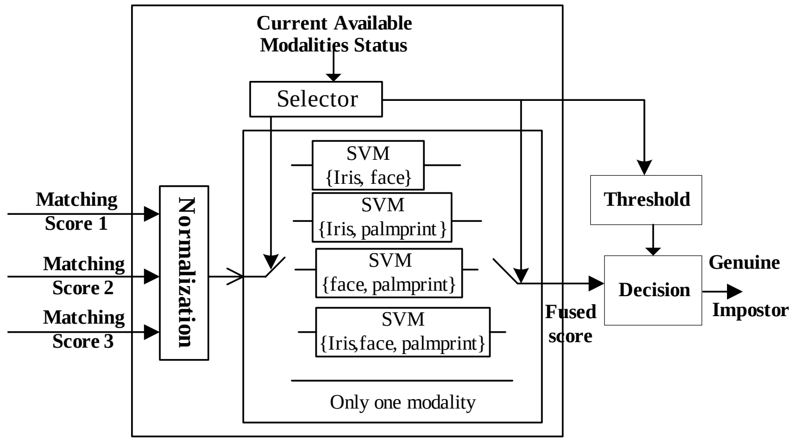


Figure 5. Structure of multiple SVMs fusion strategy

As shown in Figure 5, following the normalization of the matching scores, multiple SVMs based fusion is utilized. Multiple SVMs are parallel and each SVM is for a possible subset of {face, iris, palmprint} that contains two or three modalities. At one time, only an appropriate SVM is selected by selector according to the current available modalities status. If only one modality is available, the input score is directly as the fused score. The selected SVM can generate a corresponding fused score for the final decision. The decision threshold is corresponding to the selected SVM. At the stage of decision, if the fused score value is lower than the threshold, the applicant is accepted as a genuine user (enrolled). However, if it exceedes the threshold, the applicant is rejected as an imposter (not enrolled).

Multiple SVMs are trained using a set of training data which can be made up of the following sets: {iris, face}, {face, palmprint}, {iris, palmprint}, {iris, face, palmprint}. One set is for one SVM to learn. The principle of learning is described as follows:

SVM is based on the principle of structural risk minimization[11][12]. In this work, we use SVM to build a fusion function which can provide a fused score. Let the normalized matching scores, provided from the training sets, be combined into a multimodal score vector x. The design of a SVM trained fusion scheme consists in the estimation of a function $f: R^2(R^3) \to R$ to maximize the separability of genuine $\{f(\mathrm{x})|\text{genuine attempt}\}$ and impostor $\{f(\mathrm{x})|\text{impostor attempt}\}$ score distributions.

Suppose that the training set is $X = (\mathrm{x}_i, y_i)_{i=1}^N$, where N is the number of multimodal score vectors in the training set, and $y_i \in \{-1,1\} =$ {impostor, genuine}.

Via $\Phi: R^2(R^3) \rightarrow Q$, X is mapped into a high dimension features space *Q*. The principle of SVM relies on a linear separation in the space *Q*. In order to achieve a good level of generalization capability, the margin between the separator hyperplane $H(w, b) = \{h \in Q \mid \langle w, h \rangle_Q + b = 0\}$ and the mapped data $\Phi(X)$ is maximized (where $\langle w, h \rangle_Q$ denotes inner product in space *Q*, and $(\mathrm{w} \in Q, w_0 \in R)$ are the parameters of the hyperplane). Following the obtainment of the optimal hyperplane $H(\mathrm{w}^*, b^*)$, the decision function *D* that classifies a test pattern x_T is:

$$D(\mathrm{x}_T) = sign\{< \mathrm{w}^*, \Phi(\mathrm{x}^T) >_Q + b^*\} \qquad (4)$$

Defining $K(\mathrm{x}_i, \mathrm{x}_j) = \langle \Phi(\mathrm{x}_i), \Phi(\mathrm{x}_j) \rangle_Q$ as the kernel function (4) can be changed as:

$$D(\mathrm{x}_T) = sign\{\sum_{i \in SV} a_i y_i K(\mathrm{x}_i, \mathrm{x}_T) + b^*\} \qquad (5)$$

where $SV = \{i \mid a_i^* > 0\}$ indexes the set of support vectors.

The fusion strategy relies on the computation of the decision function *D*. In order to obtain not a final classifier decision but a fused score based on the proximity of the test pattern to the separating surface, a modification is proposed here. The fused score s_T of a test pattern x_T is defined as follows:

$$s_T = f(\mathrm{x}_T) = \sum_{i \in SV} a_i y_i K(\mathrm{x}_i, \mathrm{x}_T) + b^* \qquad (6)$$

As to the training of the SVM model, firstly, the kernel function should be decided. Several kernel functions have been put forward. However, in case of selecting the optimal kernel function, there is no fast method but by trial and error method. In this work, the radial basis function (RBF) is used as the basic kernel function by iterative trials. In the RBF kernel-based SVM, *C* and γ (kernel width) are two adjustable parameters, which play a crucial role in the performance of SVM. *C* is the regularization constant determining the trade-off between the empirical error and the regularized term, and γ underlies the mapping from input to feature space and consequently affects the performance. In our work, we adopt the grid based search method to obtain the optimal parameters (C, γ).

5. Experiments and results

5.1. Experimental database

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 [14], ORL face database[13] and PolyU palmprint database[15].

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 all SVMs, the remaining 32 subjects (224 records) as the test data to evaluate the performance of the trained system.

In a verification system the false acceptance rate (FAR) and the false rejection rate (FRR) are two widely used error measures. FAR and FRR are the functions of the decision threshold that can control the tradeoff between the two error rates. The performance of the verification system can be represented by the ROC (receive operating characteristic) curves, which plot probability of FAR versus probability of FRR for different values of the decision threshold. The point on the ROC defined by FAR=FRR is the EER point. Finally, the experiment results (ROC and EER) based on the test data, as well as some comparisons, are presented as follows.

5.2. Comparison with unimodal methods

The goal of the multimodal fusion is to achieve better precision and reliability of human authentication than single biometrics. In order to prove the effectivity of our proposed method, we present a comparison with the unimodal methods (only one modality used).

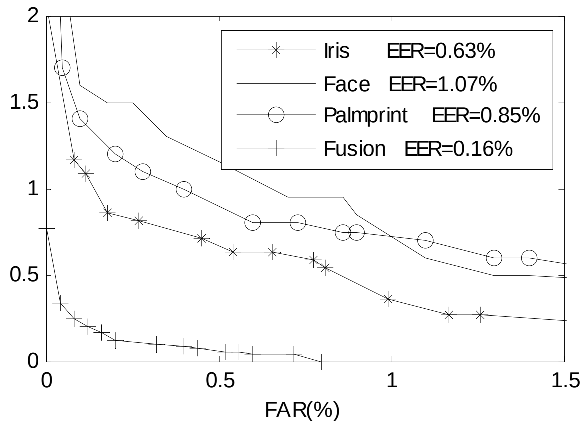


Figure 6. ROC curves of unimodal method and the proposed method

Figure 6 shows the ROC curves and EER of the following biometric systems: only iris verification, only face verification, only palmprint verification and the proposed multimodal verification in case that three modalities are all available. Iris verification is based on the improved phase information algorithm using multi-scale 2D Log-Gabor filtering, which has been described in above section and presents better performance than some current iris recognition algorithms in our previous studies[10]. Face and palmprint verifications are based on Laplacianface and Laplacian-palmprint. As can be seen from Figure 6, iris recognition usually has very high verification performance. Although UBIRIS is a noisy database and many noisy iris images are contained in the testing set, it also can achieve the performance of 0.63% EER. Face and palmprint recognition are less reliable than iris. But when three biometrics are combined using our proposed method, we can achieve a performance of 0.16% EER. This brings obvious performance improvement compared with the unimodal biometric methods. This means that multimodal biometric method is an effective way to improve human identification accuracy.

In terms of identification speed, there is no doubt that the multimodal methods will spend more time than the unimodal methods. In our experiments, we measured the average computational time required to execute the different authentication methods. All times were recorded on a Pentium-4 3.06GHz processor, running the Matlab code. As to the unimodal methods, the time for accomplishing identification with a stored template is 386 ms, 156 ms and 217 ms when iris, face and palmprint are respectively selected. As to our proposed multimodal method, when three modalities are all available and serially processed, the time for accomplishing identification is 793 ms, which is also fast enough for real-time identification. From the computational times, we can find that the time of the multimodal method is a little more than the sum of three modalities, and that the fusion part only spends very little time once SVM fusion rule has been trained in advance. Furthermore, in order to achieve the less computational time we can complement the multimodal method in the following ways: firstly, with the quick development of microelectronics, faster processors can been used to speed up the identification system; secondly, the verification modules of different modalities can been executed simultaneously by utilizing multiple processors technology especially in the embedded system, which can save more computational time.

5.3. Comparison with the previous studies

In the proposed scheme, the fusion of scores is viewed as a combination problem. Instead of some non-parametric learning fusion strategies such as sum, product and Fisher applied in previous studies [4][5], the SVM-based score level fusion strategy is employed to generate a fused score for the final decision. In the experiments, we compared the proposed SVM-based fusion strategy with the non-parametric learning fusion strategies. The detailed comparison results are as follows.

In case that three modalities are all available, Figure 7 gives the ROC curves for the mutimodal biometric methods with different fusion strategies: sum, product, Fisher and SVM. These strateges 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 performance improvement compared with unimodal method, SVM based score level fusion rule can get the best accuracy and the most improvement among four fusion strategies, which proves the superiority of combination approach based on parametric learning fusion strategies.

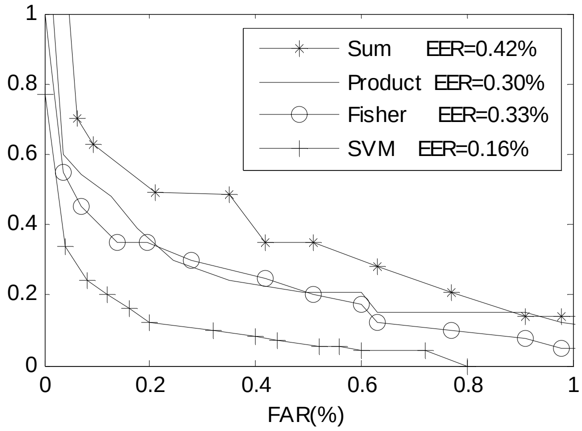


Figure 7. ROC curves of different fusion rules

In our proposed multimodal scheme, multiple parallel SVMs are utilized to overcome the limitation brought by the possible missing modalities. Next, we give the experimental results of the multiple SVMs fusion strategy when a missing modality appears.

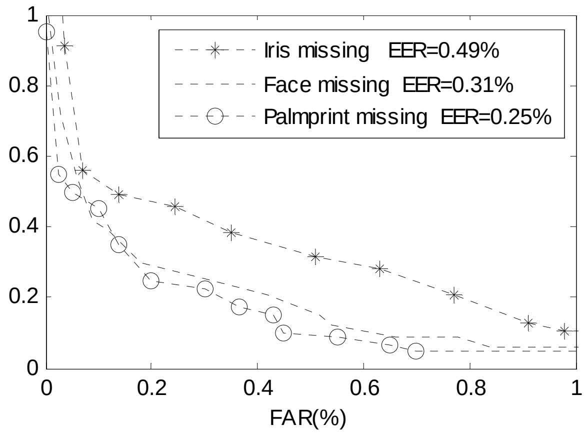


Figure 8. ROC curves of previous methods

Figure 8 represents the ROC curves of the proposed multimodal system when iris, face and palmprint are respectively missed. From Figure 8 as well as Figure 6, we can find that, although the performance of three cases is a little worse than the case that three modalities are all available, better performance also can be achieved than single biometrics system. This means that multiple SVMs fusion strategy can effectively overcome the limitation of the missing modalities.

In contrast, when the single static SVM fusion strategy is applied, the multimodal system is highly sensitive to missing modalities and the accuracy decreases noticeably. In our experiments, when iris, face and palmprint are respectively missed, the EER of the multimodal system is increased to 16.7%, 24.5% and 19.8%, which clearly shows that our multimodal authentication based multiple parallel SVMs are more practical.

6. Conclusions

In this paper, a robust multimodal biometric authentication method integrating iris, face and palmprint is proposed. Fusion of three modalities is carried out at the matching score level. Addressing the limitations of existing fusion techniques, multiple parallel SVMs fusion strategy is employed, in which all possible modality combination cases are considered and each case has a corresponding SVM to combine the scores to generate a fused score for the final decision. From the experiment results, we can conclude that:

1) Fusion of multiple biometrics can improve the verification performance comparing to the single biometrics.
2) Viewing fusion of multiple scores as a combination problem is a more flexible solution and parametric learning fusion strategy based on SVM is better than non-parametric learning fusion strategies such as sum, product and Fisher.
3) Multiple parallel SVMs fusion strategies can effectively overcome the limitation brought by the possible missing modalities. Further, the addition of a new modality does not affect the existing SVMs, instead, we can simply train the additional SVMs to handle the new modality combination, which can increase the continuity and flexibility of system.

Future work will involve investigation of better alternative verification techniques suitable for fusion of three modalities, as well as fusion of iris, face and palmprint feature at an earlier stage.

Acknowledgements

This work is funded by research grants from the National Natural Science Foundation of China (No. 60502021) and the Ph.D. Programs Foundation of Ministry of Education of China (No. 20050698025).

References

[1] **A.K. Jain, A. Ross, S. Prabhakar.** An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems 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] **Z. Liu, S. Sarkar.** Outdoor recognition at a distance by fusing gait and face. *Image and Vision Computing, Vol.*25, 2007, 817-832.

[4] **A. Ross, A. Jain.** Information fusion in biometrics. *Pattern Recognition Letters, Vol.*24, 2003, 2115-2125.

[5] **Y. Wang, T. Tan, A. K. Jain.** Combining face and iris biometrics for identity verification. *Proceedings of the* 4*th International Conference on Audio-and Video-Based Biometric Person Authentication (AVBPA), June* 9-11, 2003, *Guildford, UK, Lecture Notes in Computer Science*, 2003, *Vol.*2688, 805-813.

[6] **N.A. Fox, R. Gross, J.F. Cohn, R.B. Reilly.** Robust Biometric Person Identification Using Automatic Classifier Fusion of Speech, Mouth, and Face Experts. *IEEE Transactions on multimedia*, *Vol.*9, 2007, 701-714.

[7] **S. Ben-Yacoub, Y. Abdeljaoued, E. Mayoraz.** Fusion of Face and Speech Data for Person Identity Verification. *IEEE Transactions on Neural Networks, Vol.*10, 1999,1065-1074.

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

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

[10] **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.

[11] **B. Schölkopf, A.J. Smola.** Learning with Kernels. *MIT Press, Cambridge, MA,* 2002.

[12] **V.N. Vapnik.** The Nature of Statistical Learning Theory. *Springer*, 2000.

[13] AT&T Laboratories Cambridge. The ORL Database of Faces. *http://www.cam-orl.co.uk/facedatabase.html*.

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

[15] PolyU Palmprint. *Database: http://www.comp.polyu.edu.hk/biometrics*.

Received July 2008.