OPTO-ELECTRONICS REVIEW **17**(1), 59–64

DOI: 10.2478/s11772-008-0054-8

Multimodal biometric authentication based on score level fusion using support vector machine

F. WANG* and J. HAN

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

Fusion of multiple biometrics for human authentication performance improvement has received considerable attention. This paper presents a novel multimodal biometric authentication method integrating face and iris based on score level fusion. For score level fusion, support vector machine (SVM) based fusion rule is applied to combine two matching scores, respectively from Laplacianface based face verifier and phase information based iris verifier, 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 unimodal biometric identification methods and the previous fused face-iris methods.

Keywords: face, iris, score level fusion, support vector machine.

1. 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 behavioural characteristics to authenticate a person's identity [1].

Biometric systems based on a single source of information (unimodal systems) usually suffer from limitations like the lack of uniqueness, non-universality and noisy data [2], and hence, may not be able to achieve the desired performance requirement in real-world applications. In contrast, multimodal biometric systems combine information from multiple modalities to arrive at a decision. Some studies [2–6] have demonstrated that multimodal biometric systems can achieve better performance comparing to the individual unimodal systems.

In multimodal biometric systems, fusion of two modalities are considered a better choose because of lower cost and less complexity in real-world applications. Among all biometric technologies, face identification and iris identification both have achieved the dramatic development and promotion, and they will have the very new market behaviour in the following years. Furthermore, face and iris have many similar characters, for example, they have similar image acquisition device, and they are both non-invasive and relatively friendly. Their special 0physiological characteristics (eye is a part of face information) also indicate that the multimodal biometric systems integrating iris and face are attractive and promising. In previous studies, several integration schemes about fusion of face and iris have been developed [5,6]. In Ref. 5, Wang proposed a multimodal identification scheme based on RBF (radial basis function) neural network fusion rules. In Ref. 6, Chen applied wavelet probabilistic neural network classifier for combination of face and iris. As to recognition techniques concerned in above schemes, some earlier algorithms are used. For example, Eigenface based face recognition and local key variations based iris recognition are used in Wang's scheme, and the features of face and iris are extracted using 1D energy signal and 1D wavelet transform respectively in Chen's scheme. Recently, although some progress have been made in face recognition and iris recognition [7–9], relatively little work has been done on investigating suitable techniques from these new achievements to improve the fusion of face and iris. So in this paper we do some work in this aspect and present a fused face-iris multimodal authentication method based on score level fusion.

In our proposed fusion scheme, some new recognition techniques about iris and face are applied. As to iris verifier, an improved phase information algorithm based on 2D Log-Gabor filtering is adopted to obtain the matching score of the input iris data. And Laplacianfaces algorithm is extended by associating with the Euclidean distance to compute the matching score of the input face data. At the stage of fusion, instead of the conventional fusion rules [4], a novel fusion rule based on support vector machine (SVM) is applied to generate the fused score for the final decision. ORL face image database and UBIRIS iris image database are chosen to construct a multimodal database as the testing databases to prove the superiority of our scheme.

2. Overview of the proposed scheme

Face recognition and iris recognition both involve image preprocessing, feature extraction, matching and decision

*e-mail: fenghuawang3@gmail.com

making. Multimodal fusion for face and iris 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.

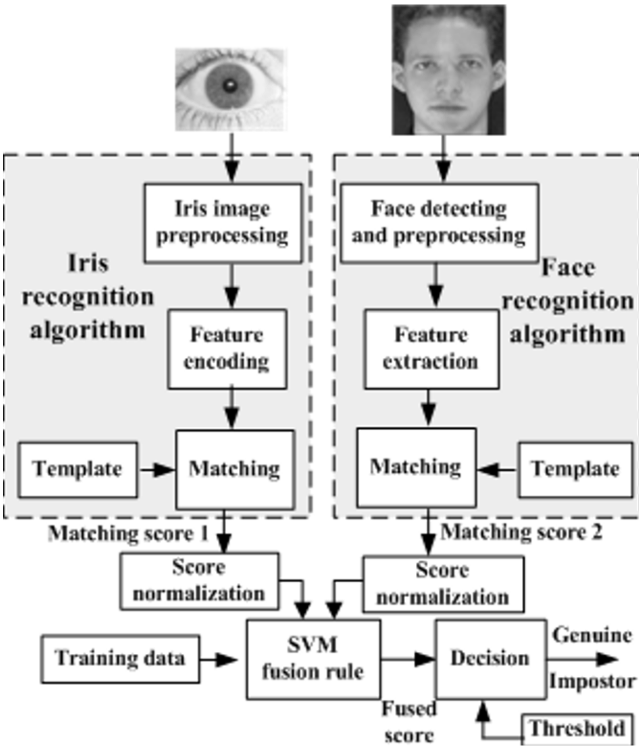


Fig. 1. Block diagram of the proposed scheme.

Figure 1 shows the block diagram of the proposed multimodal biometric authentication method integrating iris verifier and face verifier. It can be seen from Fig. 1, that face and iris images of a certain person waiting for being authenticated are first acquired. Then, the input face data are processed using Laplacianface based algorithm and the input iris data are processed using an improved phase information algorithm, respectively. Before fusion at the matching score level, the matching scores from two modalities are normalized to 0transform into a common domain because the scores generated from different modalities are heterogeneous. At the fusion stage, the normalized scores are combined using the SVM-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.

3. Face recognition and iris recognition

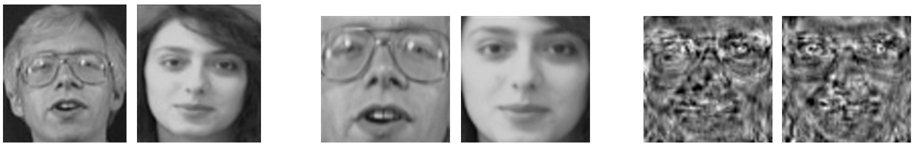
3.1. Face recognition

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 [10]. In our multimodal biometric system, the Laplacianface algorithm is employed in the face verifier part, which is a novel appearance-based face recognition algorithm [7]. Moreover, the algorithm is extended by using the Euclidean distance to compute the matching score.

Eigenface and Fisherface are two well-studied appearance-based face recognition techniques [11]. Eigenface method aims to preserve the global structure of the image space, and Fisherface method aims to preserve the discriminating information. However, the local manifold structure is more important than the global Euclidean structure in many real world classification problems especially when nearest-neighbour like classifiers are used for classification, and locality preserving projections (LPP) has discriminating power in this aspect [7,12]. So, in the Laplacianface algorithm, face images are mapped into a face subspace for analysis by using 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 Eigenface and Fisherface methods [7].

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

- 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. Figure 2(b) shows the preprocessed images,
- 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. The examples of laplacianfaces are in Fig. 2(c).
- 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



(a)
original faces

(b)
preprocessed faces

(c)
laplacianfaces

Fig. 2. Images involved in face recognition.

based on the classification experiments on the training set of the database,

- 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

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

in which, v and u are the feature vectors of matching faces. k is the dimensionality of feature vector. Following the above stage, the matching score of a face verifier is obtained as the Euclidean distance.

3.2. 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 [13] is considered a very effective one, which used Gabor filters to extract phase structure information of iris. Recent research developments [8], as well as our previous work [9], show that better performance can be achieved by using 2D Log-Gabor filters to extract phase information. 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 [9]. The detailed process is as follows:

- iris image preprocessing – prior to feature extraction, the iris image needs to be preprocessed to eliminate uninterested information and enhance interested information. The main preprocessing steps, as illustrated in Fig. 3, consist of localization of the inner and outer iris boundaries, localization of eyelid boundaries, transformation from polar coordinates to a fixed size rectangular image, and image enhancement,
- feature extraction and encoding – complex 2D Log-Gabor filters are employed to extract the phase information of iris. Similar to Ref. 13, 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. At last, the feature of iris can be described by a certain binary iris codes.
- matching – the difference between two iris images was measured by their hamming distance. Hamming distance is implemented by the simple Boolean Exclusive-OR operation (XOR) applied to the binary iris codes that encode any two iris patterns [9,14], and both of their corresponding mask bit vectors are used to prevent non-iris artifacts from influencing iris comparisons. The hamming distance can be expressed as follows

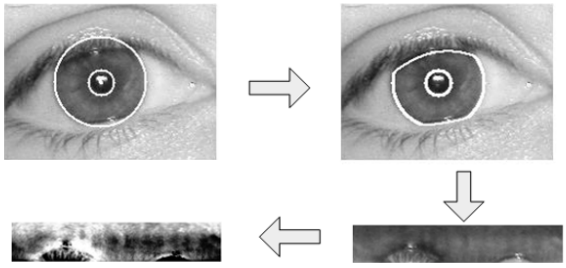


Fig. 3. Steps involved in iris preprocessing.

$$d_H = \frac{\Sigma[(codeA \otimes codeB) \cap (maskA \cap maskB)]}{\Sigma(maskA \cap maskB)}, \tag{2}$$

where *codeA* and *codeB* denote two matching iris codes, $\otimes$ denotes the Boolean Exclusive-OR operator which detects disagreement between any corresponding pair of bits, *codeA* and *codeB* denote two iris matching masks in which "0" for the non-iris regions and "1" for the iris regions, $\cap$ denotes the AND operator, which ensures that the compared bits are both deemed to have been uncorrupted by eyelids or other noise. In this part, the matching score of iris verifier is obtained as the hamming distance.

4. Score normalization

The matching scores generated from face verifier and iris 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 the matching scores d, the normalized score is given by

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

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 Eq. (3), the scores can be mapped to the [0,1] range.

5. Fusion and decision

After score normalization, the mutimodal score vector $[x_1, x_2]$ can be constructed, with x_1 and x_2 corresponding to the normalized scores of face verifier and iris verifier from a certain person waiting for being authenticated, respectively. The next step is fusion at the matching score level. This step can be approached in two distinct ways. In the first approach, the fusion is viewed as a classification problem, while in the second approach it is viewed as a combination problem. In the classification approach, the score vector is classified into one of two classes, "Accept" (genuine user) or "Reject" (impostor). In the combination approach, the score vector is combined to generate a single scalar score which is then used to make the final decision. Compared with the classification approach, the combination approach has more flexibility and can meet demands under more circumstances by adjusting the decision threshold. So in this work the combination approach based on support vector machine (SVM) is used.

SVM is based on the principle of structural risk minimization [15]. In our proposed mutimodal biometric method, we use SVM to build a fusion function which can provide a fused score.

Let the matching scores, provided by the two modalities, be combined into the multimodal score vector $x=[x_1, x_2]^T$, The design of a SVM trained fusion scheme consists in estimation of the function $f: R^2 \to R$ to maximize the separability of genuine $\{f(x)$ |genuine attempt$\}$ and impostor $\{f(x)$ |impostor attempt$\}$ score distributions.

As to the training of the SVM model, firstly, the kernel function should be decided. Several kernel functions have been put forward, but there has not been a theoretical method but usually choose it by trial and error method in case of selecting the best kernel function. 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 this study, we adopt the grid based search method to obtain the optimal parameters. To do this, we divide the training data into two sets. One of them is used to train a model, while the other, called the validation set, is used to evaluate the model. Then we set C and γ to an N*M parameter combination, which are used for SVM test. The parameter combination with the best SVM performance is chosen. Subsequently, we decrease the grid granularity, and divide the above obtained optimal parameter combination into an N*M parameter combination for further optimization, until the termination condition (the performance changes little) is satisfied. After the optimal C and γ are found, the whole training data is trained again to generate the final SVM model. In the above process, the sequential minimal optimization (SMO) training algorithm is used [16]. The advantage of SMO is that it could achieve a faster training by avoiding using the time-consuming numerical quadratic programming optimization as an inner loop.

Following the obtainment of the SVM-based fusion function, the fused score s_T of the multimodal test pattern x_T can be expressed as follows

$$s_T = f(x_T) = \sum_{i \in SV} \alpha_i^* y_i K(x_i, x_j) + w_0^*. \qquad (4)$$

Where $K(x_i, x_j)$ is the kernel function, α_i^*, w_0^* and y_i are the trained parameters. The fused score s_T can get better separability than the unimodal scores. So following its obtainment, the decision on whether impostor or genuine can be made by the predefined decision threshold which can be adjusted to reach different working points.

6. Experiments and results

To evaluate the performance of our proposed multimodal authentication method, a database containing face and iris samples is required. In this work, we construct a multimodal biometric database for our experiments by using ORL face database [17] and UBIRIS iris database [18].

The ORL data set consists of 400 frontal faces from 40 subjects (10 images of each subject). UBIRSI is a noisy database which is often used for performance evaluation of iris recognition. In our experiments, the face images from ORL are divided into two sets, 160 face images (4 images of each subject) are selected as the training samples for face recognition algorithm to create laplacianfaces, the remaining 240 face images from ORL are used to construct the multimodal database with 240 iris images from UBIRIS (40 subjects, 6 images of each subject). A face image together with an iris image is called a record, and a record is considered as an impostor attempt or a genuine attempt from a subject. So in the constructed mutimodal database, 240 records from 40 subjects (6 records of each subject) are consisted of.

In the experimental process, 48 records from 8 subjects are firstly selected as training data to estimate the parameters of SVM. The remaining 32 subjects (192 records) are as the test data to evaluate the performance of the trained system. Whether in training set or in test set, each record makes a match with all the other record in the same set using the proposed mutimodal method. Thus, in a training process, 120 intra-class matching scores (client) and 1008 inter-class matching scores (impostor) are used, and 480 intra-class matching scores and 17856 inter-class matching scores can be yielded in a testing process.

In a classification mode, Fig. 4 shows the classification results of the trained SVM with RBF kernel function in scores plane. According to Fig. 4, we can find that decision boundary by SVM with RBF kernel function can achieve better classification performance than the decision bound-

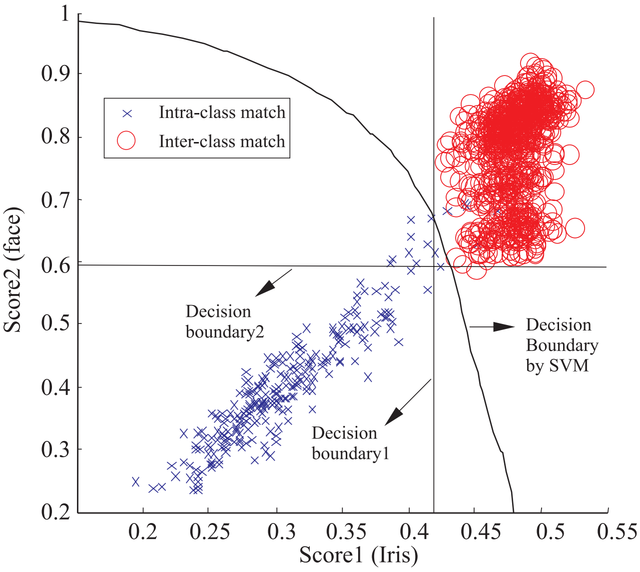


Fig. 4. Decision boundary by SVM with RBF kernel function.

aries only using single scores (Decision boundary1 and Decision boundary2 in Fig. 4).

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 plots 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.

6.1. 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 (face only and iris only).

Figure 5 shows the ROC curves and EER of the following biometric system: only iris verification, only face verification, and the proposed multimodal verification. Iris verification is based on a 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 [9]. Face verification is based on Laplacianface algorithm. As it can be seen from Fig. 5, iris recognition usually has very high verification performance, although many noisy iris images are contained in the testing database, it also can achieve the performance of 1.06% EER. Face recognition is less reliable than iris. But when two biometrics are combined using our proposed method, we can

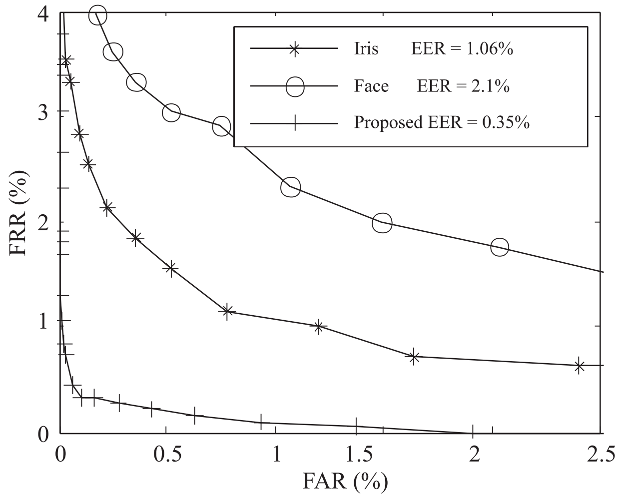


Fig. 5. ROC curves of unimodal method and the proposed method.

achieve a performance of 0.35% EER, and brings obvious performance improvement compared with two unimodal biometric methods. This means that multimodal biometric method is an effective way to improve human identification accuracy. Moreover, the multimodal verification methods also increase the difficulty of imposters' faking the biometric.

6.2. Comparison with the methods using conventional score level fusion rules

In the proposed method, the SVM-based score level fusion rule has been employed. To evaluate the performance of different fusion rules, we also tested some conventional score level fusion rules [4] (such as sum, product and fisher) when fusing at the score level. Then, we compared these methods using the conventional fusion rules with our proposed method using SVM-based fusion rule. The detailed comparison results are as follows.

Figure 6 gives the ROC curves for the mutimodal biometric methods with different fusion rules, sum, product,

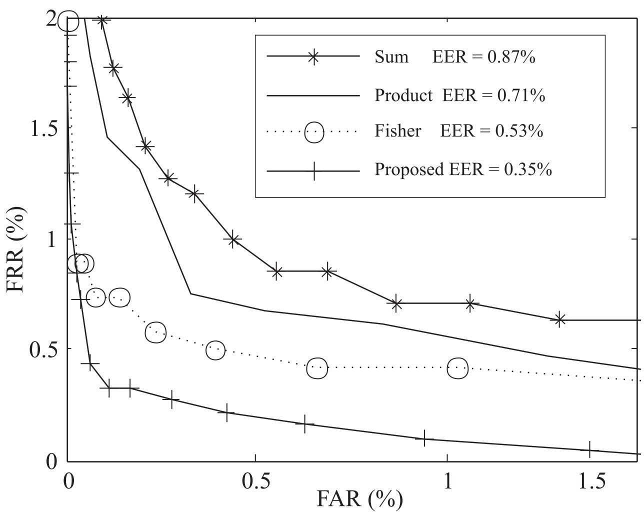


Fig. 6. ROC curves of different fusion rules.

fisher, and SVM. From the figure, we can see that although these fusion rules 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 rules, which prove the superiority of SVM based rule.

6.3. Comparison with the previous studies

Some integration schemes about fusion of face and iris have been proposed in the previous literatures [5,6]. In this section, we make a comparison with Wang's method [5] and Chen's method [6], the details about which has been described in the introduction section. The following figure represents the ROC curves and EER of our method and the two previous methods.

As shown in Fig. 7, our method also achieves the best performance compared with Wang's method and Chen's method. We think the better performance of our method mainly comes from two reasons, one is the application of new recognition techniques to get the matching scores of face and iris, and the other is the application of SVM-based fusion rule. Moreover, SVM-based fusion rule not only brings better recognition performance but also reduces the training time and computational cost comparing to neural network based fusion rule applied in Wang's method and Chen's method.

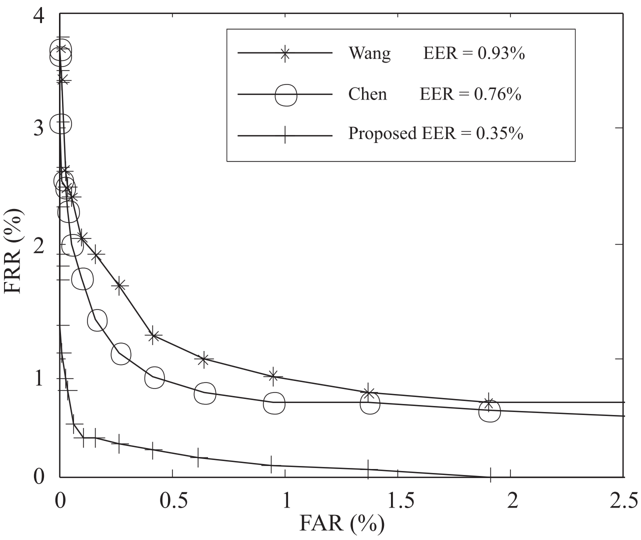


Fig. 7. ROC curves of previous methods.

7. Conclusions

In this paper, a multimodal biometric identification method integrating face and iris based on score level fusion was proposed. For score level fusion, two matching scores respectively from face verifier and iris verifier was combined by using SVM based fusion rule to generate a single scalar score which is used to make the final decision. From the experiment results, we can conclude that:

- fusion of the two biometrics can improve the verification accuracy than the single biometrics,
- SVM-based fusion rule can achieve better fusion effect than the conventional score level fusion rules such as sum, product and fisher *et al*.,
- the proposed method have the superiority over the previous methods due to the application of the new recognition algorithms and SVM-based fusion rule.

Future work will involve investigation of better alternative recognition technique suitable for fusion of face and iris, as well as fusion of face and iris feature at an earlier stage.

References

1. A.K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition", *IEEE T. Circ. Syst. Vid.* **14**, 4–20 (2004).
2. A.K. Jain and A. Ross, "Multibiometric systems", *Commun. ACM* **47**, 34–40 (2004).
3. Z. Liu and S. Sarkar, "Outdoor recognition at a distance by fusing gait and face", *Image Vision Comput.* **25**, 817–832 (2007).
4. A. Ross and A.K. Jain, "Information fusion in biometrics", *Pattern Recogn. Lett.* **24**, 2115–2125 (2003).
5. Y. Wang, T. Tan, Y. Wang, and D. Zhang, "Combining face and iris biometric for identity verification", *Proc. 4th Int. Conf. on Audio- and Video-Based Biometric Person Authentication* (*AVBPA*) **1**, 805–813 (2003).
6. C. Chen and C. Chu, "Fusion of face and iris features for multimodal biometrics", *Lect. Notes Comput. Sc.* **3832**, 571–580 (2006).
7. H. Xiaofei, Y. Shuicheng, H. Yuxiao, P. Niyogi, and Z. Hong-Jiang, "Face recognition using Laplacianfaces", *IEEE T. Pattern Anal.* **27**, 328–340 (2005).
8. Y. Du, "Using 2D log-gabor spatial filters for iris recognition", *Biometric Technology for Human Identification* **6202**, 1–8 (2006).
9. F. Wang and J. Han, "Iris recognition method using Log-Gabor filtering and feature fusion", *J. Xi'an Jiaotong Univ.* **41**, 889–893 (2007).
10. M. Yang, D.J. Kriegman, and N. Ahuja, "Detecting faces in images: A survey", *IEEE T. Pattern Anal.* **24**, 34–58(2002)
11. P.N. Belhumeur, J.P. Hepanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection", *IEEE T. Pattern Anal.* **19**, 711–720 (1997).
12. D. Cai, X. He, J. Han, and H.J. Zhang, "Orthogonal Laplacianfaces for face recognition", *IEEE T. Image Process.* **15**, 3608–3614 (2006).
13. J. Daugman, "How iris recognition works", *IEEE T. Circ. Syst. Vid.* **14**, 21–30(2004).
14. J. Daugman, "The importance of being random: Statistical principles of iris recognition", *Lect. Notes Comput. Sc.* **36**, 279–291 (2003).
15. B. Scholkopf and A.J. Smola, *Learning with Kernels*, MIT Press, Cambridge, MA, 2002.
16. S. Keerthi, S.K. Shevade, and C. Bhattacharyya, "Improvements to Platt's SMO algorithm for SVM classifier design", *Neural Computation* **13**, 637–649 (2001).
17. AT&T Laboratories Cambridge, *The ORL Database of Faces*: http://www.cam-orl.co.uk/facedatabase. html.
18. H. Proenca and A. Alexandre, *UBIRIS Iris Image Database*: http://iris.di.ubi.pt.

