Chapter 9

Multimodal Biometrics Based on Near-Infrared Face Recognition

Rui Wang, Shengcai Liao, Zhen Lei, and Stan Z. Li

9.1 INTRODUCTION

Biometric identification makes use of the physiological or behavioral characteristics of people, such as fingerprint, iris, face, palmprint, gait, and voice, for personal identification [1], which provides advantages over nonbiometric methods such as password, PIN, and ID cards. Its promising applications as well as the theoretical challenges have gotten its heated attraction from the last decade.

Face recognition is a natural, nonintrusive and easy way for biometrics and has been one of the most popular techniques. However, most current face recognition systems are based on face images captured in visible light spectrum, which are compromised in accuracy by changes in environmental illumination. The near-infrared (NIR) face image-based recognition method [2–4] overcomes this problem. It is shown to be invariant to the changes of the visible lighting and hence is accurate and robust for face recognition.

Recent research [5–9] has pointed out that multimodal biometric fusion can significantly improve the performance of the system due to the complementary information from different modalities helpful for classification. There exists various methods for multimodal fusion. Brunelli and Falavigna [10] proposed a person identification system based on voice and face, using a HyperBF network as the best performing

Biometrics: Theory, Methods, and Applications. Edited by Boulgouris, Plataniotis, and Micheli-Tzanakou

fusion module. Kittler et al. [11] proposed a face and voice multimodal biometric system and developed a common theoretical framework for combing classifiers in reference 5 with several fusion techniques including sum, product, minimum, and maximum rules, where the best combination results are obtained for a simple sum rule. Hong and Jain [6] proposed an identification system based on face and fingerprint, where fingerprint matching is applied after pruning the database via face matching. Ross et al. [12] combined face, fingerprint, and hand geometry biometrics with sum, decision tree, and linear discriminant-based methods, where the sum rule achieves the best. Wang et al. [13], Son and Lee [14], and Chen and Chu [15] developed face and iris multimodal biometric systems, and different fusion methods were investigated. Kumar et al. [16] described a hand-based verification system that combined the geometric features of the hand with palmprints at the feature and matching score levels. Li et al. [9] proposed a systematic framework for fusing 2D and 3D face recognition at both feature and score levels, by exploring synergies of the two modalities at these levels and achieved good performance in large database. Chang et al. [7] combined ear and face biometrics with an appearance-based method. Ribaric and Fratric [8] described a biometric identification system based on eigenpalm and eigenfinger features, with fusion applied at the matching score level.

In this chapter we present a near-infrared (NIR) face-based approach for multimodal biometric fusion. The motivations for this approach are the following: (1) NIR face recognition overcomes problems arising from uncontrolled illumination in visible light (VL) image-based face biometric and achieves significantly better results than VL face; and (2) the fusion of NIR face with VL face or iris biometrics is a natural way for multibiometrics, because it is either face-based (NIR face + VL face) or NIR-based (NIR face+ irais).

The NIR face is fused with VL face or iris modality at the matching score level. As for score level fusion, there are two common approaches. One is to treat it as a combination problem, in which the individual matching scores are combined according to some rule such as sum rule, max rule, or min rule to generate a single scalar score. The others is to formulate it as a classification problem, such as LDA [12] or a power series model (PSM)-based method [17]. The latter needs to be learned in a training set.

We evaluate these fusion methods on real large multimodal databases we collected, in which NIR face and VL face or iris image for one subject are captured simultaneously by our own image capture device. The NIR database is publicly available on the web [18]. The experimental results show that the learning-based fusion methods such as LDA and PSM are comparatively better than other conventional methods.

The rest of this chapter is organized as follows: Section 9.2 briefly introduces the near-infrared face recognition and describes the fusion of NIR face with VL face and the fusion of NIR face with iris modality respectively. Section 9.3 describes several fusion methods. The experimental results and discussions are presented in Section 9.4, and in Section 9.5 we conclude the chapter.

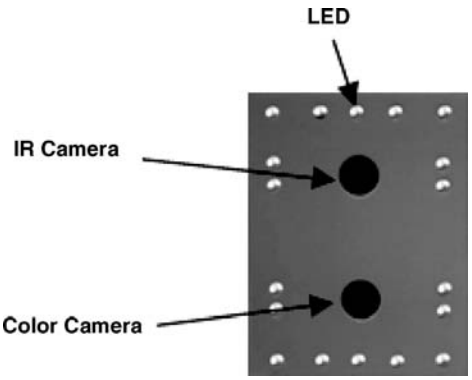

Figure 9.1. The capture device of NIR face images.

9.2 NIR FACE-BASED MULTIBIOMETRICS

9.2.1 NIR Face Recognition

The NIR face image is taken using the device shown in Figure 9.1, composed of NIR LEDs, an NIR camera, and a color camera. The NIR LED lights are approximately coaxial to the lens direction. For NIR camera, to minimize ambient lights in visible spectrum, a long-pass optical filter can be used with the lens to cut off visible light while allowing NIR light to pass. We choose a filter such that ray passing rates are 0%, 50%, 88%, and 99% at the wavelength points of 720, 800, 850, and 880 nm, respectively. The filter cuts off visible environmental lights ($<$ 750 nm) while allowing 80–90% of the 850-nm NIR light to pass. Such NIR image is captured in a good condition regardless of visible lights in the environment and encodes intrinsic information of the face, subject only to a monotonic transform in the gray tone [2]. Based on this, we use local binary pattern (LBP) features further to compensate for the monotonic transform, thus deriving an illumination invariant face representation for face recognition.

LBP is introduced as a powerful local descriptor for microfeatures of images [19]. The LBP operator labels the pixels of an image by thresholding the 3×3 neighborhood of each pixel with the center value and considering the result as a binary number (or called LBP codes). An illustration of the basic LBP operator is shown in Figure 9.2. Note that the binary LBP code is circular.

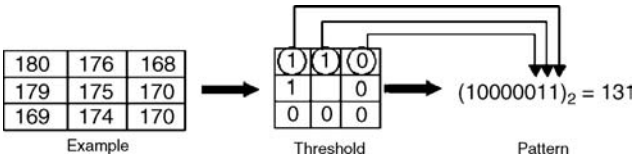

Figure 9.2. Calculation of LBP code from 3×3 subwindow.

Input: Sequence of N weighted examples:
$\{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$;
Initialize: $w_i = \frac{1}{N}$, $i = 1, 2, ..., N$, $F(x) = 0$
Integer T specifying number of iterations;
For $t = 1, \ldots, T$

(a) Fit the regression function $f_t(x)$ by weighted least squares of y_i to x_i with weights w_i.
(b) Update $F(x) \leftarrow F(x) + f_t(x)$
(c) Update $w_i \leftarrow w_i e^{-y_i f_t(x_i)}$ and renormalize.

Output: the final classifier, $sign[F(x)] = sign[\sum_{t=1}^{T} f_t(x)]$

Figure 9.3. Gentle AdaBoost algorithm in reference 19.

LBP histograms over local regions provides a more reliable description when the pattern is subject to alignment errors. Hence, in our work a histogram of the base LBP codes is computed over a local region centered at each pixel, and it is considered as a set of individual features. The original LBP feature pool is of high dimensionality. Not all of them are useful or equally useful, and some of them may cause a negative effect on the performance. Therefore, we adopt the following AdaBoost algorithm [19] to select the most discriminative and complementary features and construct the powerful classifier for face recognition.

AdaBoost iteratively learns a sequence of weak classifier $f_t(x)$ and linearly combines them to construct a strong classifier $F(x)$. At each iteration, a weak classifier $f_t(x)$ is chosen to minimize the weighted squared error $J_{wse} = \sum_{i=1}^{N} w_i(y_i - f_t(x_i))^2$.

Biometric recognition is a multiclass problem, whereas the above AdaBoost learning is for two classes. To deal with this problem, we take the similar measure in reference 20 to construct intrapersonal and extrapersonal classes to convert the multiclass problem into a two-class one. Here, instead of deriving the intra- or extrapersonal variations using difference images as in reference 20, the training examples for our learning algorithm is the set of differences between each pair of LBP histogram features at the corresponding locations. The positive examples are derived from pairs of intrapersonal and the negative from pairs of extrapersonal differences.

With the two-class scheme, the face matching procedure will work in the following way: It takes the probe face image and a gallery face image as the input, computes a difference-based feature vector from the two images, and then calculates a similarity score for the feature vector using some matching function. A decision is made based on the score, to classify the feature vector into the positive class (coming from the same person) or the negative class (different persons).

9.2.2 Fusion with VL Face

It may be advantageous to combine information contained in different face modalities to overcome limitations in single face data so as to improve the performance

of system. Several methods have been proposed to combine information from multiface modalities to achieve higher performance. Heo and co-workers [21] present two approaches to fuse thermal infrared (TIR) and visible light (VL) face images. One is to average image pixels of the two modalities, and the other is to fuse them at the decision level. In references 22 and 23, the 2D and 3D face information is dimensionally reduced, then a classifier is built and the scores are fused by sum rule. Pan and coworkers 24 capture 31 different multispectrum face images to be fused for recognition and obtain a good result. However, the image capture devices of TIR face images, 3D face images, and multispectrum face images are all complex and expensive, which is disadvantageous for practical application. The proposed modalities, NIR and VL face images, due to its complementary information and the low-cost image acquisition, may be a good choice for the fusion.

Figure 9.4 illustrates the block diagram of the fusion framework of NIR and VL multimodal faces at the score level. The input is a pair of NIR and VL face

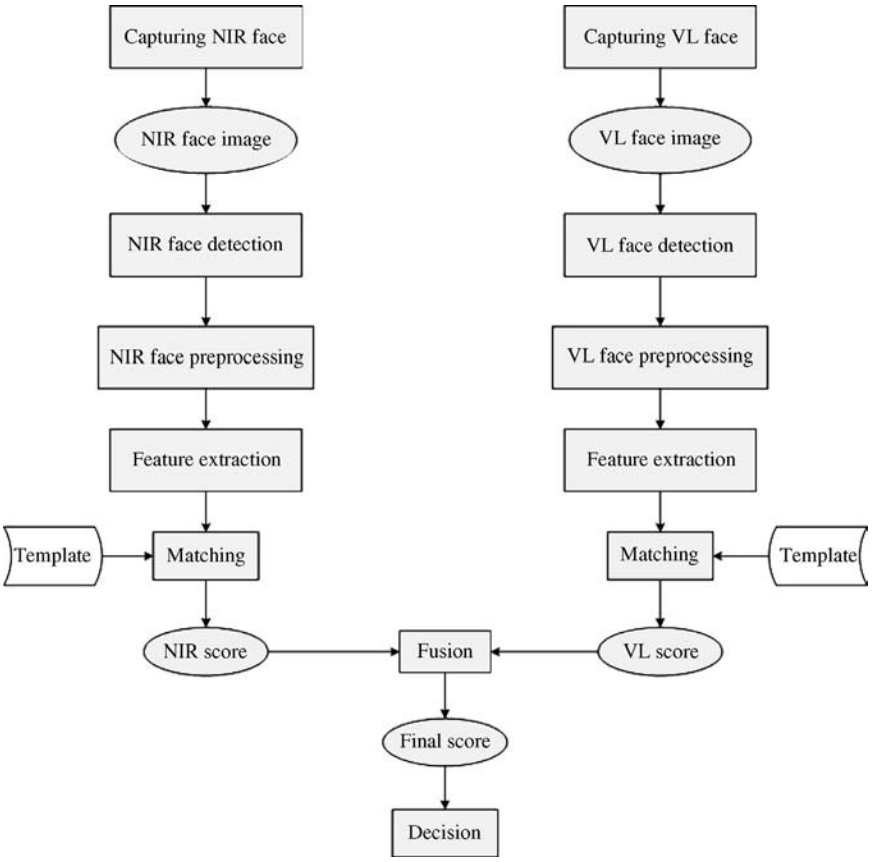

Figure 9.4. Algorithm structure for score fusion of NIR face and VL face.

images. After face and eye detection processes, the NIR and VL faces are cropped and normalized to a pre-fixed size. The LBP features for NIR and VL are then extracted and fed into the NIR and VL face recognition engines (e.g., AdaBoost classifiers), respectively, to produce two scores for NIR and VL. Finally, the two matching score are fused according to some rule and compared to a threshold to make the final decision.

In this paper, VL face recognition uses the same recognition algorithm as NIR, in which AdaBoost learning [25] is used to construct a powerful classifier based on a Local Binary Pattern (LBP) representation [19].

9.2.3 Fusion with Iris

Fusing NIR face and iris modality is another choice for NIR face-based multimodal biometrics, and it brings the following advantages [26]. (1) The NIR face and iris images can be acquired simultaneously by an improved commercial digital camera. (2) The NIR face and iris contain different or complementary information for recognition, so that the total error rate (the sum of false accept rate and false reject rate) is known to go down [27]. (3) It reduces spoof attacks on the biometric system because of the difficulty in making fake iris images.

For iris recognition, we adopt the well-known algorithm of Daugman [28], which includes the following four steps. (1) It is necessary to localize the inner and outer boundaries of the iris precisely and to detect and exclude the eyelids if they intrude. (2) The portion of the image corresponding to the iris is translated into a normalized form, so that possible dilation of the pupil does not affect the system. (3) The feature extraction process is completed by the use of 2D Gabor wavelets to perform a multiscale analysis of the iris. The information about local phase, coded with two bits corresponding to the signs of the real and imaginary parts, is obtained, which is a 256-byte IrisCode. (4) Similarity scores are obtained by computing a hamming distance between two IrisCodes.

In this chapter, NIR face and iris modalities are acquired using a single high-resolution camera with active frontal NIR lighting. This not only is a natural way for face and iris multimodal biometrics, since both NIR face and iris need active NIR nodality, but also brings convenience to the user. Figure 9.5 summarizes the structure of the algorithms of NIR face and iris biometrics fusion using a single high-resolution NIR face image. The input is a high-resolution NIR face image. The face and eyes are localized accurately using a face and eye detection algorithm [2]. After that, the left and right irises are segmented from the face, and both the face and irises are normalized into a pre-fixed sizes. The facial LBP features and iris Gabor features are extracted and fed into NIR face and iris recognition engines respectively to be compared to the corresponding templates. Finally, the three matching scores are fused following some rule, and they are compared with a threshold to make the final classification.

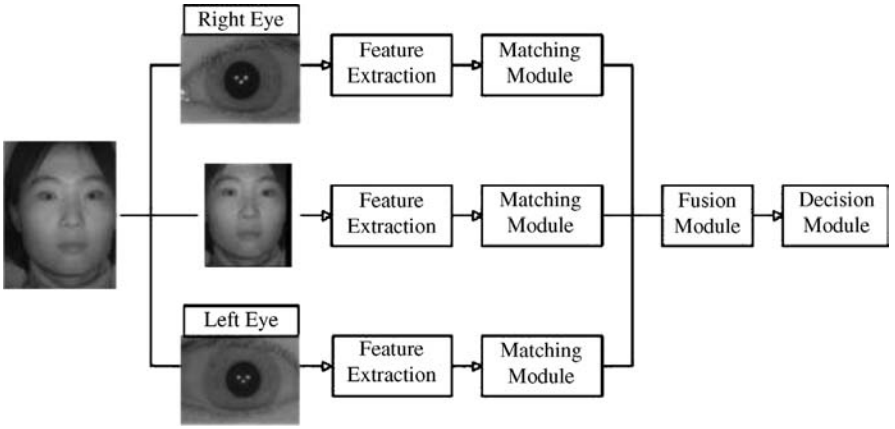

Figure 9.5. Algorithm structure for score fusion of face, left iris, and right iris.

9.3 METHOD OF MULTIBIOMETRICS FUSION

The NIR face biometric is fused with VL face or iris modality at score level. There are two common approaches for fusion at the matching score level: (1) Treat it as a combination problem, in which the individual matching scores are combined follow some rule such as sum rule, max rule, min rule, and so on, to generate a single scalar score and the other is to formulate it as a classification problem, such as the LDA- and PSM-based method.

Suppose we have M scores from M biometric modalities corresponding to one sample pair. The sum rule outputs the summation of the M scores. The min rule outputs the minimum value of the M scores, and the max rule outputs the maximum value of the M ones.

For the LDA-based method, the M scores are formulated as a M-dimension vector. The sample pairs are divided into two classes of intra and extra which denote the samples from the same persons and different persons, respectively. The purpose of LDA is to find an optimal projective direction that maximizes the between-class scatter while it minimizes the within-class scatter. The LDA-based method can be essentially considered as a weight sum rule.

Recently, Toh [17] proposed a power series model (PSM)-based fusion method. If we denote the scores from M biometric modalities of one sample pair as $\bar{s} = (s_1, s_2, \ldots, s_M)$, where $s_k, k = 1, \ldots, M$, is the kth score corresponding to the kth modality, then a power series model is constructed as

$$f(\alpha, \bar{s}) = \alpha_0 + \sum_{r=1}^{R} \sum_{m=1}^{M} \alpha_{r,m} s_m^r, \tag{9.1}$$

where $f(\alpha, \bar{s})$ is the fusion score, α_0 and $\alpha_{r,m}$ are the model parameters to be determined, and R is the order of PSM. There are totally $K = 1 + RM$ parameters. The

above power series model can be rewritten by matrix formulation

$$f(\bar{\alpha}, \bar{s}) = \sum_{k=0}^{K-1} \alpha_k p_k(\bar{s}) = \bar{p}(\bar{s}) \cdot \bar{\alpha}, \tag{9.2}$$

where $\bar{p}(\bar{s}) = [p_1(\bar{s}), p_2(\bar{s}), \ldots, p_K(\bar{s})]$ and $\bar{p}_k(x)$ is a row vector corresponds to a power basis expansion term. $\bar{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_K]^T$ denotes the parameter vector to be estimated.

Specifically, if $R = 1$, Eq. (9.1) equals

$$f(\bar{\alpha}, \bar{s}) = \alpha_0 + \sum_{m=1}^{M} \alpha_m s_m. \tag{9.3}$$

As α_0 is a constant, the PSM can be considered as the weight sum rule. Moreover, if we have $\alpha_m = 1, \alpha_0 = 0$, the PSM will be degenerated into sum rule.

We can use least square to estimate the parameter $\bar{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_k]^T$, by minimizing the loss function:

$$J(\bar{\alpha}) = \frac{1}{2} \|\bar{y} - P\bar{\alpha}\|_2^2, \tag{9.4}$$

where $\| \bullet \|_2$ denotes the Euclidean distance,

$$P = \begin{bmatrix} p_1(\bar{s}_1) & p_2(\bar{s}_1) & \cdots & p_K(\bar{s}_1) \\ p_1(\bar{s}_2) & p_2(\bar{s}_2) & \cdots & p_K(\bar{s}_2) \\ \vdots & \vdots & \ddots & \vdots \\ p_1(\bar{s}_n) & p_2(\bar{s}_n) & \cdots & p_K(\bar{s}_n) \end{bmatrix}$$

and $\bar{y} = [y_1, y_2, \ldots, y_n]^T$, where $y_i \in \{1, 0\}$ is the class label which denotes the sample pair from the same person or different persons.

The solution of Eq. (9.4) can be obtained as

$$\bar{\alpha} = (P^T P)^{-1} P^T \bar{y} \tag{9.5}$$

Given a testing sample $\bar{s}_t$ as input, we can get

$$f_t = f(\bar{\alpha}, \bar{s}_t) = \bar{p}(\bar{s}_t) \cdot \bar{\alpha}, \tag{9.6}$$

where $\bar{\alpha}$ is computed in Eq. (9.5). Assuming a threshold τ, the matching result of $\bar{s}_t$ can then be determined as

$$\begin{cases} \bar{s}_t \in w_1 & \text{if } f_t \geq \tau, \\ \bar{s}_t \in w_0 & \text{if } f_t < \tau. \end{cases} \tag{9.7}$$

Figure 9.6. Typical VL face examples (upper) and NIR face examples (lower) in database.

9.4 EXPERIMENTS

9.4.1 Databases

To evaluate the performance of the proposed multimodal fusion in real-world applications, real multimodal databases are built for two fusions.

(a) Database for NIR Face and VL Face. For the fusion of NIR and VL face, the capture device consists of two CMOS cameras. One is for the NIR image and the other is for the VL image, with resolution of 640*480 (in pixel). Therefore, a pair of NIR and VL face are captured from one object simultaneously.

All the face images are taken near-frontal but in an uncontrolled indoor environment with varying pose, expression, and lighting. Some examples of typical NIR and VL face pairs in the database are shown in Figure 9.6. Both NIR and VL face images are then cropped into 144×112 according to the eye coordinates detected automatically. Figure 9.7 shows some examples of the cropped images.

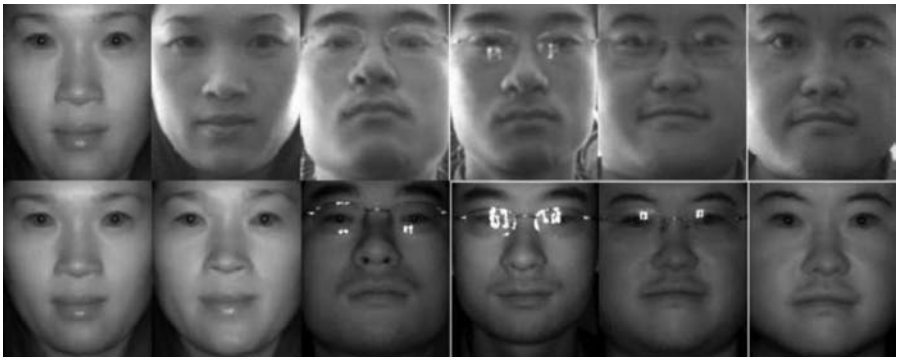

Figure 9.7. Cropped VL face examples (upper) and NIR face examples (lower) in database.

The NIR and VL face database is composed of 3940 pairs of images from 197 subjects, with 20 pairs per person. All the images are divided into training set and test set randomly. The training set includes 3000 pairs of images from 150 subjects, while the test set includes the left 940 pairs of images from 47 subjects. So the training set and the testing set have no intersection of persons and images either. In the training phase, we construct the AdaBoost classifiers for NIR and VL face modalities respectively and utilize the training set for LDA and PSM learning based fusion. In testing phase, each input NIR face and VL face image pair is matched with all of the other image pairs in the test set. This generates $47 \times C_{20}^2 = 8930$ intraclass (positive) and $20 \times 20 \times C_{47}^2 = 432,400$ extra-class (negative) samples.

(b) Database for NIR Face and Iris. To capture a high-resolution image including face and iris information sufficiently, we use a 10-megapixel CCD digital camera with up to 3648×2736 pixels. The camera is placed about 60–80 cm away from the subject. Around the camera lens, active NIR LED lights of 850 nm are mounted to provide frontal lighting. We use a band-pass optical filter on the camera lens to cut off visible light while allowing NIR light to pass.

An NIR face + iris database is built containing 560 high-resolution (2736*3648 pixels) NIR images. It includes 112 subjects of 55 females and 57 males, aged from 17 to 35, with 10 images for 76 subjects and 5 images for other 34 subjects. Figure 9.8

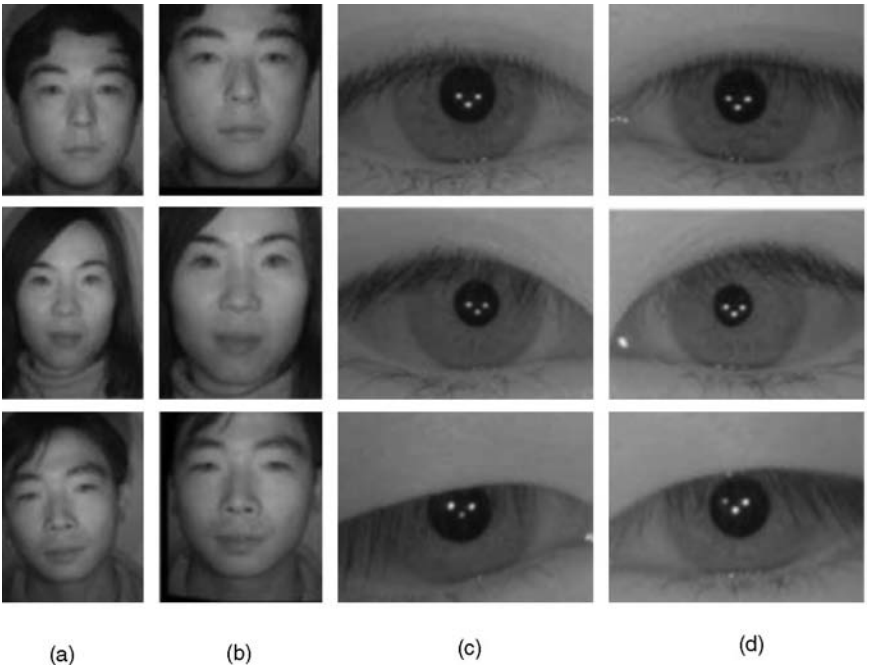

Figure 9.8. A high-revolution face image and separated face and both iris images. **(a)** High resolution NIR face images. **(b)** NIR image segmented from (a). **(c)** Left iris segmented from (a). **(d)** Right iris segmented from (a).

Table 9.1. Relationship Between GAR and the Order of the Power Series Model on Training Data

R	1	2	3	4	5	6	7
GAR(%) (FAR = 0.1%)	95.6	94.9	95.7	95.9	94.8	95.4	95.7
EER(%)	1.21	1.35	1.17	1.11	1.33	1.20	1.16

shows some examples of face images and the segmented iris parts. The training set includes 250 images from 50 subjects. The test set includes 310 images from 62 subjects, which are totally different from the subjects of the training set.

9.4.2 Results

9.4.2.1 Results for NIR Face and VL Face Fusion

For PSM-based fusion, the parameter order R influences the performance of the fusion algorithm, so it needs to be optimized first. To determine the value of parameter R, we use the training set to evaluate the performance of varying the value of parameter. Table 9.1 shows the genuine acceptance rate (GAR) when the false acceptance rate (FAR) is at 0.1% and the equal error rate (EER) is at various values of R.

From the result, we can see that the PSM-based fusion method achieves the lowest error rate when R is 4 in the training set. Therefore, in the following experiments, we choose $R = 4$ for the PSM-based method.

In this experiment, AdaBoost classifier is used in both NIR and VL face recognition. The output score of AdaBoost is a posterior probability $P(y = +1|x)$ that ranges from 0 to 1. Thus both output scores of NIR and VL face classifiers are well normalized in [0, 1] by AdaBoost, and no further score normalization process is needed when fusing them. We compare six score-level fusion methods: PSM [17], LDA [29], sum rule, product rule, min rule, and max rule [30]. Table 9.2 shows the match results of

Table 9.2. GAR and EER for Score Fusion of NIR Face and VL Face

	GAR(%) (FAR = 0.1%)	EER(%)
PSMSF	93.2	1.84
LDA	92.2	1.94
SUM	91.7	2.80
PRODUCT	91.7	2.81
MIN	89.4	4.56
MAX	90.7	2.04
NIR	90.1	2.34
VIL	84.0	5.27

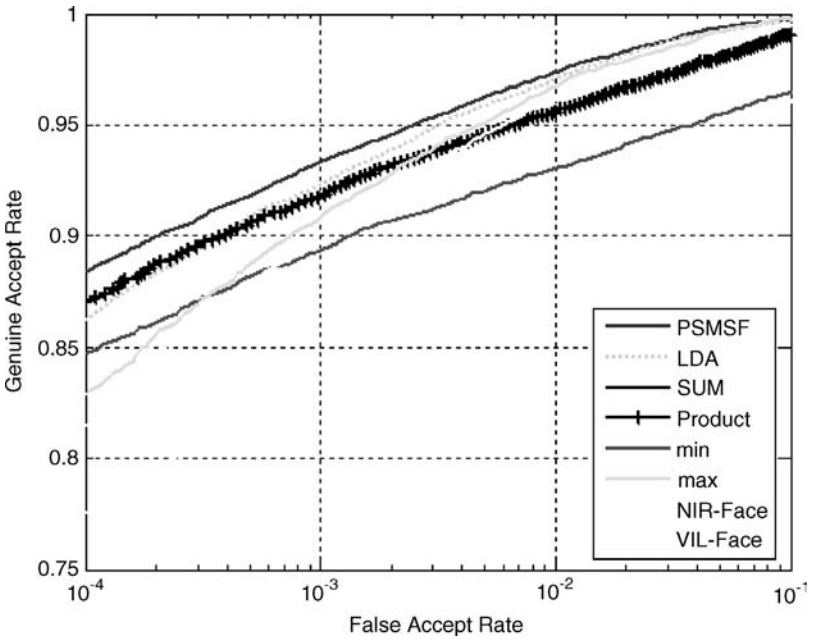

Figure 9.9. ROC curves for score fusion of NIR face and VL face.

six fusion methods and two single modalities, and Figure 9.9 shows the corresponding ROC curves.

9.4.2.2 Results for NIR Face and Iris Fusion

In this section we use the AdaBoost classifier trained from the above experiment for NIR face recognition and construct iris classifier using the method in reference 31. The method to choose R is the same as fusion of NIR and VL face and the determined value of R is 3 for the NIR face and iris fusion.

Since the face matching scores and iris matching scores are not in common domain, we need to normalize the scores from different modalities first. Three common normalization methods—min-max, Z-score, and tanh-score normalization—are used and compared in the experiments. Table 9.3 shows the match results for six fusion methods and three single modalities with three different normalization methods, and Figure 9.9 shows the corresponding ROC curves with different normalization methods.

9.4.3 Discussions

From the experimental results, we can observe that in most cases, fusion of NIR with VL faces and fusion of NIR face with iris modality can improve the recognition

Table 9.3. GAR and EER for Score Fusion of NIR Face and Iris

	GAR(%) (FAR = 0.1%)			EER(%)		
	Min–Max	*Z*-Score	Tanh	Min–Max	*Z*-Score	Tanh
PSM	98.9	98.9	98.9	0.39	0.39	0.39
LDA	98.6	98.6	98.4	0.44	0.52	0.53
SUM	98.3	97.8	97.8	0.67	1.19	1.20
MIN	88.8	97.7	97.7	1.61	0.46	0.46
MAX	91.4	92.0	92.0	5.34	5.52	5.53
NIR face	88.2	88.2	88.2	1.59	1.59	1.59
Left-iris	91.1	91.1	91.1	3.99	3.99	3.99
Right-iris	91.8	91.8	91.8	5.08	5.08	5.08

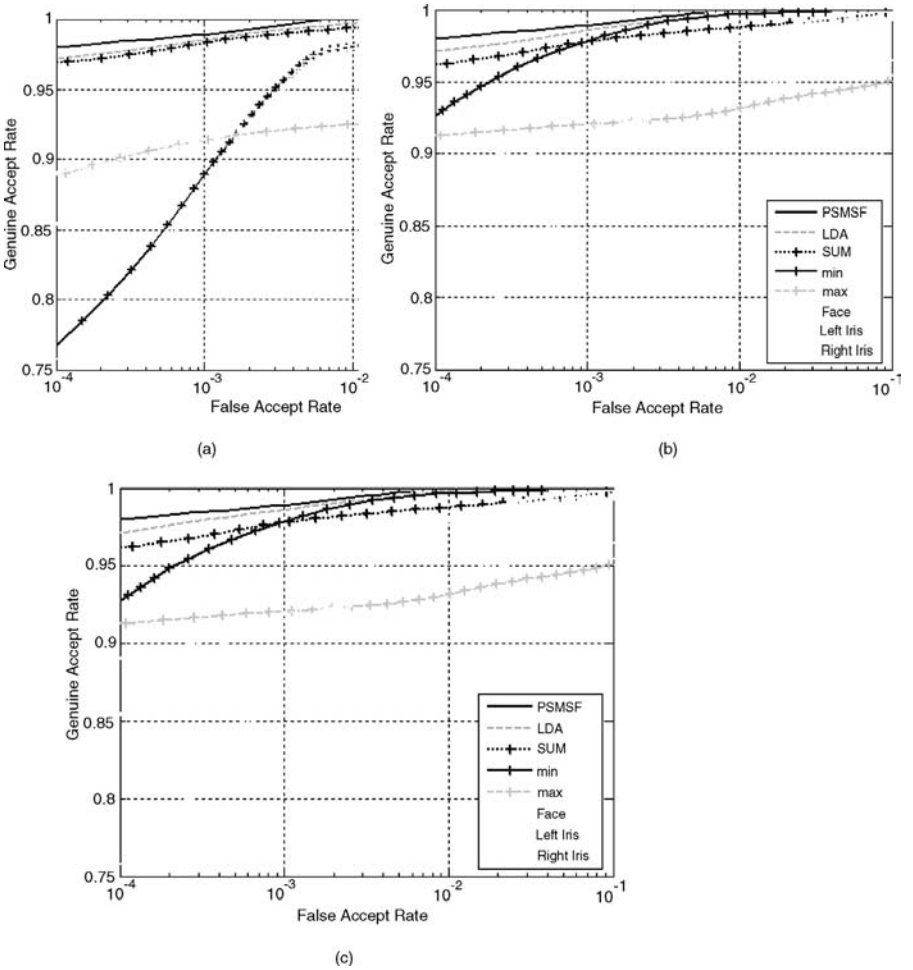

Figure 9.10. ROC curves for score fusion of NIR face and iris with three normalization methods: (**(a)** min–max, **(b)** *Z*-score, **(c)** Tanh.)

accuracy compared to any single modality performance, which proves the effectiveness of multimodal biometrics. The learning-based fusion methods such as LDA and PSM achieve better results than other methods, and the PSM-based method has achieved the best performance in all the cases. In the case of fusion of NIR and VL faces, the genuine accept rate (GAR) increases from 90.1% (NIR face) to 93.2%; and in the case of fusion of NIR face and iris biometric, the GAR increases from 88.2% (NIR face) to 98.9%. Moreover, comparing the results in NIR face and iris fusion, it can be seen that the PSM- and LDA-based methods have similar results with different score normalization methods, while the performance of some of the conventional methods such as min or max rule may fluctuate a little large. This indicates that the learning-based PSM and LDA methods are more robust to normalization methods and hence more suitable in practical applications.

9.5 CONCLUSIONS

In this chapter we explore synergies of NIR + VL faces and NIR face + iris by proposing an NIR face-based approach for multibiometrics. The NIR face is fused with VL face or iris in a natural way. This approach takes the advantages of recent progress in NIR face recognition, and it further improves the performance of biometric systems. Experimental results show that both the fusion of NIR + VL face and NIR face + iris can significantly improve the system performance in real databases. The learning-based methods such as LDA- and PSM-based fusion achieve the best results and are robust to score normalization, thus they are practical in real applications.

ACKNOWLEDGMENTS

This work was supported by the following funding resources: National Natural Science Foundation Project #60518002, National Science and Technology Supporting Platform Project #2006BAK08B06, National 863 Program Projects #2006AA01Z192 and #2006AA01Z193, Chinese Academy of Sciences 100 people project, and the AuthenMetric Collaboration Foundation.

REFERENCES

1. A. K. Jain, R. M. Bolle, and S. Pankanti, *Biometrics: Personal Identification in Networked Society*, Kluwer, Norwell, MA, 1999.
2. S. Z. Li, R. Chu, S. C. Liao, and L. Zhang, Illumination invariant face recognition using near-infrared images, *IEEE Trans. Pattern Anal. Mach. Intell.* **29**(4):627–639, 2007.
3. S. Z. Li, L. Zhang, S. C. Liao, X. X. Zhu, R. F. Chu, M. Ao, and R. He. A near-infrared image based face recognition system, in *Proceedings of 7th IEEE International Conference Automatic Face and Gesture Recognition (FG-2006)*, Southampton, UK, April 10–12, 2006, pp. 455–460.
4. S. Z. Li, R. F. Chu, M. Ao, L. Zhang, and R. He, Highly accurate and fast face recognition using near infrared images, in *Proceedings of IAPR International Conference on Biometric (ICB-2006)*, Hong Kong, January 2006, pp. 151–158.
5. J. Kittler, M. Hatel, R. P. W. Duin, and J. Matas, On combining classifiers, *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(3):226–239, 1998.

6. L. Hong and A. K. Jain, Integrating faces and fingerprints for personal identification, *IEEE Trans. Pattern Anal. Mach. Intell.* **20**:1295–1307, 1998.
7. K. Chang, K. Bowyer, V. Barnabas, and S. Sarkar, Comparison and combination of ear and face images in appearance based biometrics, *IEEE Trans. Pattern Anal. Mach. Intell.*, **25**:1160–1165, 2003.
8. S. Ribaric and I. Fratric, A Biometric identification system based on eigenpalm and eigenfinger features, *IEEE Trans. Pattern Anal. Mach. Intell.* **27**:1698–1709, 2005.
9. S. Z. Li, C. S. Zhao, X. X. Zhu, and Z. Lei, Learning to fuse 3d+2d based face recognition at both feature and decision levels, in *Proceedings of IEEE International Workshop on Analysis and Modeling of Faces and Gestures*, Beijing, China, October 16, 2005, 44–54.
10. R. Brunelli and D. Falavigna, Person identification using multiple cues, *IEEE Trans. Pattern Anal. Mach. Intell.*, **17**:955–966, 1995.
11. J. Kittler, G. Matas, K. Jonsson, and M. Sanchez, Combining evidence in personal identity verification systems, *Pattern Recognit. Lett.* **18**:845–852, 1997.
12. A. Ross, A. K. Jain, and J. Z. Qian, Information fusion in biometrics, *Pattern Recognit. Lett.* **24**:2115–2125, 2003.
13. Y. Wang, T. Tan, and A. K. Jain, Combining face and iris biometrics for identity verification, in *Proceedings of International Conference on Audio- and Video-Based Person Authentication*, 2003, pp. 805–813.
14. B. Son and Y. Lee, Biometric authentication system using reduced joint feature vector of iris and face, in *Proceedings of International Conference on Audio- and Video-Based Person Authentication*, 2005, pp. 513–522.
15. C. Chen and C. T. Chu, Fusion of face and iris features for multimodal biometrics, in *Proceedings of IAPR International Conference on Biometric*, January 5–7, 2006, pp. 571–580.
16. A. Kumar, D. C. M. Wong, H. C. Shen, and A. K. Jain, Personal verification using palmprint and hand geometry biometric, in *Proceedings of International Conference on Audio- and Video-based Person Authentication*, 2003, pp. 668–678.
17. K.-A. Toh, Error-rate based biometrics fusion, in *ICB*, 2007, pp. 191–200.
18. S. Z. Li et al., CBSR NIR face dataset, http://www.cse.ohio-state.edu/otcbvs-bench/.
19. A. Hadid, T. Ahonen, and M. Pietikainen, Face description with local binary patterns:application to face recognition, *IEEE Trans. Pattern Anal. Mach. Intell.* **28**:2037–2041, 2006.
20. B. Moghaddam, C. Nastar, and A. Pentland, A Bayesian similarity measure for direct image matching, Media Lab Technical Report No.393, MIT, August 1996.
21. B. Abidi, J. Heo, S. Kong, and M. Abidi, Fusion of visual and thermal signatures with eyeglass removal for robust face recognition, in *IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum in Conjunction with CVPR 2004*, 2004, pp. 94–99.
22. K. W. Bowyer K. I. Chang, and P. J. Flynn, Multi-Modal 2d and 3d biometrics for face recognition, in *IEEE Workshop on Analysis and Modeling of Faces and Gestures*, 2003.
23. K. W. Bowyer, K. I. Chang, and P. J. Flynn, Face recognition using 2D and 3D facial data, in *Workshop in Multimodal User Authentication*, 2003, pp. 25–32.
24. M. Prasad, Z. H. Pan, G. Healey, and B. Tromberg, Face recognition in hyperspectral images, *IEEE Trans. Pattern Anal. Mach. Intell*. **25**:1552–1560, 2003.
25. P. Viola and M. Jones, Robust real-time object detection, *Int. J. Comput. Vis.* **00**:000–000, 2002.
26. S. Pankanti, L. Hong, and A. Jain, Can multibiometrics improve performance?, in *Proceedings AutoID'99*, 1999, pp. 947–950.
27. A. K. Jain, Y. Wang, and T. Tan, Combining face and iris biometrics for identity verification, *AVBPA 2003 LNCS* **2688**:805–813, 2003.
28. J. G. Daugman, High confidence visual recognition of persons by a test of statistical independence, *IEEE Trans. Pattern Anal. Mach. Intell.*, **15**:1148–1161, 1993.
29. G. J. Mclachlan, *Discriminant Analysis and Statistical Pattern Recognition*, 1996.
30. R. P. W. Duin, J. Kittler, M. Hatel, and J. Matas, On combining classifiers, *IEEE Trans. Pattern Anal. Mach. Intell.* **20**(3):226–239, 1998.
31. J. Friedman, T. Hastie, and R. Tibshirani, Additive logistic regression: A statistical view of boosting, *Ann. Stat.* **28**(2):337–374, April 2000.