MVA2013 IAPR International Conference on Machine Vision Applications, May 20-23, 2013, Kyoto, JAPAN

An Efficient Feature-Level Fusion Scheme in Multimodal Biometrics

Ren-He Jeng and Wen-Shiung Chen
VIP-CCLab., Dept. of Electrical Engineering
National Chi Nan University, Taiwan
E-mail: wschen@ncnu.edu.tw

Lili Hsieh
Dept. of Information Management
Hsiuping University of Science and Technology
Taiwan

Abstract

Due to inherent properties in each biometric, none of any single biometric can guarantee 100% authentication accuracy by only itself. Information fusion is the key step in multimodal biometric recognition systems. Previous researches have confirmed that fusion at low-level considered to be effective, but difficult. In this paper, a feature-level fusion scheme based on shuffle/coding rule is proposed and three types of fusion are addressed. Some competitive experimental results show that the proposed fusion scheme is very advantageous.

1. Introduction

Previous works [1]-[6] organized multimodal biometrics into two major approaches, pre-classification and post-classification. Pre-classification refers to combining information on senor- or feature- level. The post-classification combines classification information, such as matching sources and decision.

The previous researches confirmed that fusing information at early stages will be much effective than at later stages. Since the source data and feature data contain rich information, fusing information at pre-classification is more useful than at post-classification [6]. Despite its improving accuracy, multimodal biometric has several drawbacks. First, since recognition is a serial steps, it may appear error propagation problems. Second, combining non-homogenous features is difficult and even impossible, because the relationship of each feature spaces may not be known. Third, concatenating two feature vectors result in a feature vector with huge dimensionality leading to the 'curse of dimensionality' problem. Therefore, few researchers have studied integration at the feature level. Besides, it is too rigid to combine information at post-classification. We will consider both the implementation complexity and information quantity. Therefore, we focus mainly on feature-level fusion.

2. The Proposed Feature-Level Fusion

2.1 Shuffle-based Feature-Level Fusion

In our proposed approach, feature extractors extract feature vectors by using different algorithms and quantizing scales. Since feature spaces are uncorrelated, it is difficult or even impossible to combine them. We suppose that even two distinct groups of features are not correlation, the distance of a pair of features are correlation. To describe the distances, the quantization scales should be properly considered. Therefore, our approach normalizes features in distinct independent spaces by using different normalization parameters. Since feature spaces are independent, it can reserve the relativity between original data in new feature spaces.

In this section, a fusion scheme, called shuffle-based feature-level fusion (S-FLF), is to combine two (or more) feature vectors to form a new feature via "shuffle" operation. The S-FLF consists of the following processes: (i) feature extraction, (ii) quantization, and (iii) Shuffle/Coding, as shown in Figure 1.

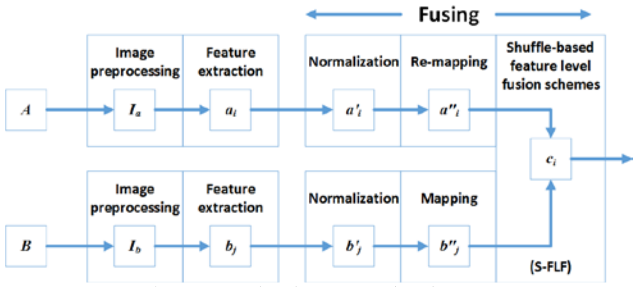


Figure 1. The framework of S-FLF.

Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be two feature vectors, obtained by any different feature extraction algorithms and defined as follows:

$$\boldsymbol{a} = [a_1, \cdots, a_i, \cdots, a_m], \; i = 1, 2, ..., m,$$

and

$$\boldsymbol{b} = [b_1, \cdots, b_j, \cdots, b_n], \; j = 1, 2, ..., n.$$

A normalization algorithm with the different parameter settings is used to normalize the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ to the vectors

$$\boldsymbol{a}' = [a'_1, \cdots, a'_i, \cdots, a'_m], \; i = 1, 2, ..., m,$$

and

$$\boldsymbol{b}' = [b'_1, \cdots, b'_j, \cdots, b'_n], \; j = 1, 2, ..., n,$$

where a'_i and b'_j are real numbers and $0 < a'_i, b'_j < 1$. Then, a re-mapping process is used to generate two new vectors

$$\boldsymbol{a}'' = [a''_1, \cdots, a''_i, \cdots, a''_m], \; i = 1, 2, ..., m,$$

and

$$\boldsymbol{b}'' = [b''_1, \cdots, b''_j, \cdots, b''_n], \; j = 1, 2, ..., n,$$

where a''_i and b''_j are real numbers and $0 < a''_i, b''_j < p$. Finally, a shuffle/coding rule generates a new binary codes

$$\boldsymbol{c}_s = [c_1, \cdots, c_i, \cdots, c_{sc}], \; i = 1, 2, ..., sc,$$

where $sc = \min\{m, n\}$ and $\boldsymbol{c}_i$ is a binary code vector of length p, denoted by

$$\boldsymbol{c}_i = (c_{i1}, \cdots, c_{ij}, \cdots, c_{ip}), \; j = 1, 2, ..., p, \text{ for each } i.$$

2.2 Quantization Process

A. Scale Normalization

Consider the feature vectors $\boldsymbol{x}_i$ of each class

$$\boldsymbol{X} = \left[\boldsymbol{x}_1, \cdots \boldsymbol{x}_i, \cdots, \boldsymbol{x}_k\right],\ i = 1, 2, ..., k\ .$$

where the feature vectors $\boldsymbol{x}_i$ has t elements.

$$\boldsymbol{x}_i = \left(x_1, ..., x_j, ...x_t\right),\ j = 1, 2, ..., t\ .$$

The feature vector $\boldsymbol{x}$ is normalized into the range of [0, 1]. In this paper, three normalization rules, such as (i) max-min, (ii) fuzzy and (iii) max-mean, are adopted as follows:

(i) Max-min: Max-min normalization is a process of taking data measured in its engineering units and transforming it to a value between 0.0 and 1.0, i.e.,

$$x_i'^j = \left(x_i^j - v_j\right) / \left(V_j - v_j\right),$$

where the minimum (V_j) and maximum (v_j) values of j-th feature are computed.

(ii) Fuzzy: Fuzzy normalization is defined as follows:

$$x_i'^j = 0.5 + 0.5 \sin\left(\pi / \left(V_j - v_j\right)\left[x_i^j - 0.5\left(V_j - v_j\right)\right]\right).$$

(iii) Max-mean: This rule only considers the positive samples:

$$x_i'^j = \left(\left(x_i^j - \mu_j - m_j\right) / \left(M_j - m_j\right) + 0.5\right) p$$

where the mean value of j-th feature (u_j) and the minimum (m) and maximum (M) values are, respectively,

$$m_j = \min\left(x_i^j - \mu_j\right)$$

and

$$M_j = \max\left(x_i^j - \mu_j\right).$$

B. Re-mapping

Next, a re-mapping process with a linearfunction is used to map x'_k into a real-value variable x''_k within the range of [0, p], as:

$$x''_k = \lfloor p x'_k \rfloor,\ k = 1, 2, ..., o\ .$$

We try to limit the influence of variables in the range of [0, 255] in experiments and change quantization parameter p. The experimental systems were tested using three different normalization algorithms for two image scales. The experimental results are good enough for p = 8, as shown in Figure 2. Therefore, we assume that the parameter p should be set to be about (upper limit of raw data interval)/32.

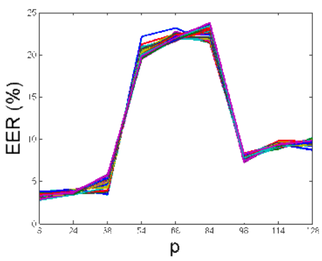


Figure 2. The effect of p to EER on our approach.

2.3 Shuffle/Coding

A shuffle/coding algorithm creates feature codes $\boldsymbol{c}_s$=[c_1,…,c_i,…,c_{sc}] for $\boldsymbol{a}''$ and $\boldsymbol{b}''$ (actually, a_i and b_i). In fact, our approach outset decided the feature selected. Each *shuffle codec*$_i$ is the length of p. We illustrate the idea of our approach by presenting the different between a_i'' and b_i'' (actually, a_i and b_i) in codes. If $a''_k \le b''_k$, this algorithm returns a binary code $\boldsymbol{c}_i$=(c_{i1},…,c_{ij},…,c_{ip}), j=1,2,…,p, for each i, in which the bit

$$\boldsymbol{c}_{ij} = \begin{cases} 1 & a''_k \le j \le b''_k \\ 0 & \text{others} \end{cases},\ i, k = 1, 2, ..., sc\ ,\ j = 1, 2, ..., p\ ,$$

that is, the code of each bit between bit-(a''_k-1) and bit-(b''_k-1) is set to be "1" and the other codes are "0." Conversely, if $a''_k > b''_k$, it returns a binary code c_i by

$$\boldsymbol{c}_{ij} = \begin{cases} 1 & j < b''_k,\ a''_k > j \\ 0 & \text{others} \end{cases},\ i, k = 1, 2, ..., sc\ ,\ j = 1, 2, ..., p\ ,$$

Theoretically, this approach can be written as:

$$\boldsymbol{c}_i = \begin{cases} \left[2^{b''_k} - 2^{\left(a''_k - 1\right)}\right]_2 & a''_k \le b''_k \\ \left\{\left(2^p - 1\right) - \left[2^{a''_k} - 2^{\left(b''_k - 1\right)}\right]\right\}_2 & a''_k > b''_k \end{cases},$$

$$k = 1, 2, ..., sc\ ,$$

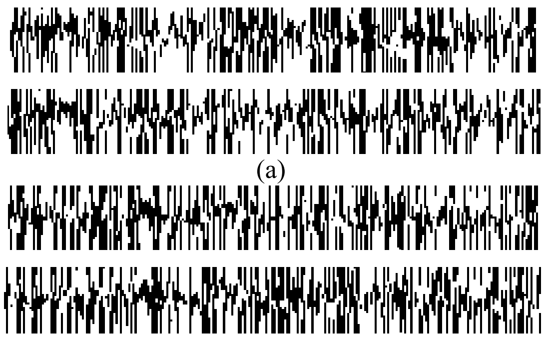
where

$$\boldsymbol{c}_i = \left(c_{i1}, \cdots, c_{ij}, \cdots, c_{ip}\right),\ \forall \boldsymbol{c}_{ij} = \{0, 1\}\ .$$

Finally, according to the re-mapping values of all feature vectors, the feature codes is

$$\boldsymbol{c}_s = \left[\boldsymbol{c}_1, ..., \boldsymbol{c}_i, ..., \boldsymbol{c}_{sc}\right],\ i = 1, 2, ..., sc\ .$$

Figure 3 shows some S-FLF codes for different classes, in which there are two samples for each class.



(a)

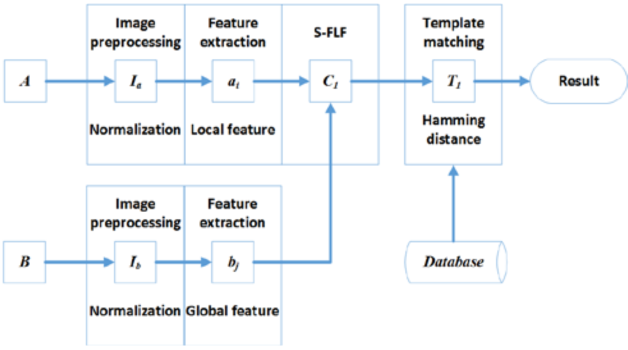
(b)

Figure 3. Some examples of shuffle/coding operation.

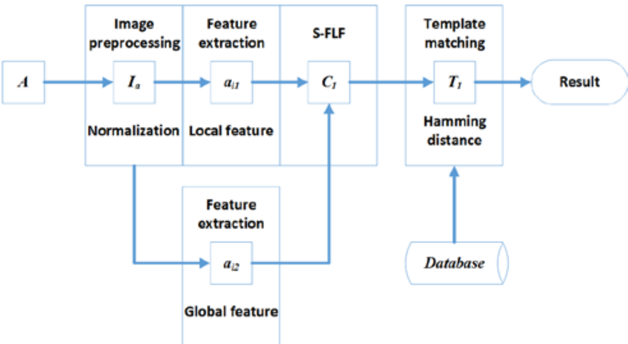
3. Experimental Results

In this paper, we combined three different biometrics, such as face [7], iris [8], and palmprint [9]. The performance was evaluated by equal error rate (EER), which is the value where the false accept rate (FAR) and false reject rate (FRR) are equal.

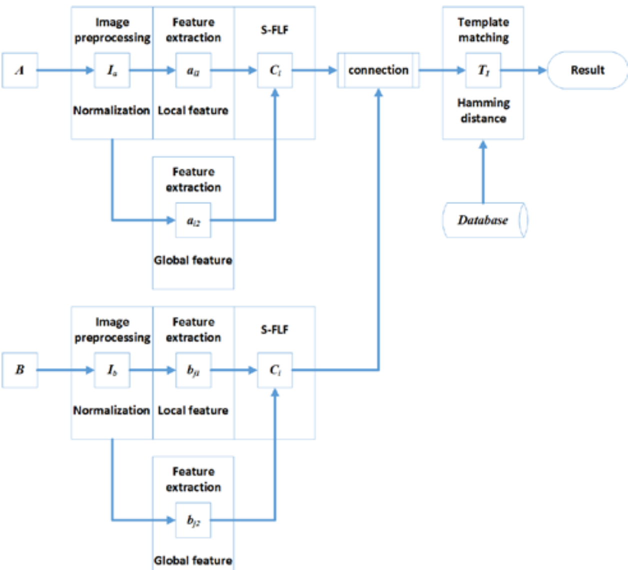
According to a variety of combinations of feature vectors for generating S-FLF codes, three different types of S-FLF fusion scheme are addressed, as shown in Figure 4. (i) Type 2: Two biometric traits were extracted features, and combined them, as shown in Figure 4(a). (ii) Type 1: A biometric trait was extracted multiple features, and combined them, as shown Figure 4(b). (iii) Type 3: The third type is simply extended from Type 2. In Type 2, the system demands users to provide their own biometrics as inputs. Each trait was encoded individually, and the system combines them, as shown in Figure 4(c).



(a) S-FLF Type 1: Multiple features were extracted for multiples traits.



(b) S-FLF Type 2: Multiple features were extracted for a trait.



(c) S-FLF Type 3: Multiple features were extracted for each trait and connected.

Figure 4.The framework of three types of S-FLF scheme.

In Table 1 and Table 2, the experiments are tested on two image scales. For scale 1, the scales includes: (1) iris: 128×32, and (2) face, hand: 64×64. Besides, the scales 2 includes (1) iris: 256×64, and (2) 128×128. The experimental results show high accuracy for (a) that combining the local features. Since experimental results (c) are better than (d), our approach combines non-homogenous features by (c).

Table 1 The accuracy of S-FLF Type 1 (p=8). H: palmprint, F: face, I: iris, left option: ***a***, right option: ***b***.

Scale 1		max-min	fuzzy	min-mean
a-local, ***b***-local.	H-F	2.2	2.3	2.2
	H-I	2.0	1.6	1.8
	F-H	2.5	2.1	2.5
	F-I	5.0	4.6	4.5
	I-H	2.1	1.8	2.0
	I-F	5.0	4.7	3.8
a-global, ***b***-global.	H-F	28.8	31.2	32.9
	H-I	15.9	17.6	18.1
	F-H	28.9	32.4	33.5
	F-I	21.6	24.5	22.9
	I-H	14.6	16.7	20.2
	I-F	19.8	22.8	23.0
a-local, ***b***-global.	H-F	4.9	4.6	5.6
	H-I	2.0	2.0	3.6
	F-H	14.4	16.4	14.8
	F-I	11.7	12.8	11.3
	I-H	7.5	9.0	7.1
	I-F	8.7	10.0	8.1
a-global, ***b***-local	H-F	12.8	15.7	15.7
	H-I	7.1	8.4	8.1
	F-H	4.7	4.8	4.9
	F-I	8.9	9.3	8.0
	I-H	1.9	1.8	3.7
	I-F	10.5	11.1	11.6

Table 2 The accuracy of S-FLF Type 1 (p=8). H: palmprint, F: face, I: iris, left option: ***a,*** right option: ***b***.

Scale 1		max-min	fuzzy	min-mean
a-local, ***b***-local.	H-F	1.7	2.3	2.1
	H-I	1.9	2.1	1.6
	F-H	1.9	2.1	2.0
	F-I	4.4	4.5	2.9
	I-H	1.9	1.9	1.6
	I-F	4.3	4.6	2.6
a-global, ***b***-global.	H-F	25.5	29.7	28.9
	H-I	20.9	22.6	22.8
	F-H	25.9	29.1	28.9
	F-I	23.7	34.2	28.2
	I-H	21.0	20.1	23.7
	I-F	23.5	34.9	29.4
a-local, ***b***-global.	H-F	9.6	9.2	12.5
	H-I	4.9	4.5	7.0
	F-H	7.5	8.6	8.3
	F-I	11.0	13.0	13.5
	I-H	7.4	7.0	5.0
	I-F	12.0	12.6	10.3
a-global, ***b***-local	H-F	7.9	7.9	9.8
	H-I	7.0	7.2	6.0
	F-H	9.5	8.1	11.7
	F-I	11.7	11.9	10.4
	I-H	4.4	4.6	8.2
	I-F	12.2	12.8	15.8

Since that feature codes are generated by encoding a pair of features, we can observe the difference of both features. That is, we compare the difference between feature-a and feature-b for all traits, and counting them. The histogram shows that the distributions are concentrated on negative, as shown in Figure 5. Therefore, the features-a should be smaller than the feature-b.

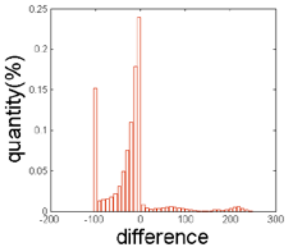


Figure 5. The distribution of the difference in features.

Tables 1~3 provide that combining information is improving the accuracy for recognizing people. Experimental results confirmed our approach and show a great potential for combining information in feature-level.

Table 3 The accuracy of S-FLF Type 2. (p=8)

	max-min	fuzzy	min-mean
H（64×64）	3.4	3.6	3.8
H（128×128）	4.5	3.7	5.2
F（80×80）	16.9	18.3	18.0
F（160×160）	18.2	17.9	20.7
I（128×32）	8.6	9.3	8.1
I（256×64）	7.4	8.3	7.2

Using the most suitable parameter of p and deciding the feature comparing model, our proposed system was completed, as listed in Table 4.

Table 4 The accuracy of S-FLF Type 3. (p=32)

		max-min	fuzzy	min-mean
Scale 1	HF	2.9	3.0	2.5
	HI	1.9	1.6	1.7
	FI	4.7	5.8	5.0
	HFI	1.3	1.0	1.4
Scale 2	HF	3.9	3.5	3.8
	HI	1.7	2.0	1.6
	FI	4.7	5.6	4.7
	HFI	1.4	1.8	1.6

Again, we measure the performance for robust of the fused recognition under two experimental scenarios: (1) system built with noisy data of one biometric and two noiseless biometrics. (2) The system built with noisy data of two biometrics and a noiseless biometric. The system adds 10-80% Gaussian noise to an image.

In Figure 6, the experimental results are tested on S-FLF Types 3 for combining Face (64×64), Hand (64×64), and Iris (32×128). Since our feature codes are dependent on their own features and created codes are independent, in Figure 6 (b), there are two robust traits for keeping high accuracy of fused system. One is palmprint and the other is iris. Therefore, multimodal biometric system is feasible for development on feature-level.

4. Conclusions

This work proposed an information fusion scheme at feature level for multimodal biometrics. Unlike previous works, our approach describes the relationship between two feature spaces. It does not need too much resource and is simple. The experimental results reveal that the proposed tactics can improve the performance further.

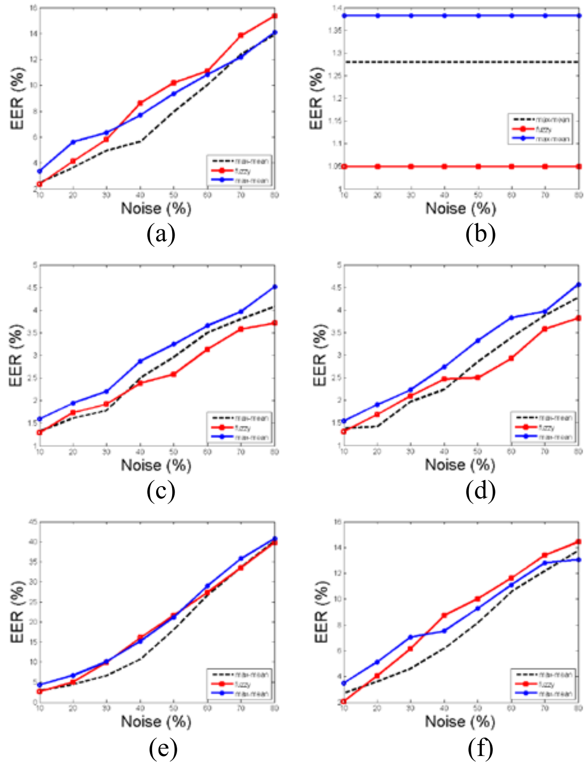


Figure 6. Experimental results with variable Gaussian noise on (a) palmprint, (b) face, (c) iris, (d) face and iris, (e) palmprint and iris, and (f) palmprint and face.

References

[1] A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of Multibiometrics*, 1st ed. Springer, 2006.

[2] B. Bhanu and V. Govindaraju, *Multibiometrics for Human Identification*, 1st ed. Cambridge University Press, 2011.

[3] R. S. Blum and Z. Liu, *Multi-sensor Image Fusion and Its Applications*, CRC Press, 2006.

[4] A. Ross1, A. Jain and J. Z. Qian, "Information fusion in biometrics," *International Conference on Audio- and Video-Based Person Authentication*, vol. 1, pp. 354-359, 2000.

[5] C. Sanderson and K. K. Paliwal, "Information fusion and person verification using speech and face information," *Research Paper IDIAP-RR 02-33*, 2002.

[6] A. K. Jain, K. Nandakumar, and A. Ross, "Score normalization in multimodal biometric systems," *Pattern Recognition*, vo. 38, no. 12, pp. 2270-2285, 2005.

[7] FEI Face Database. http://fei.edu.br/~cet/facedatabase.html.

[8] H. Proença and L. A. Alexandre, "UBIris: A noisy iris image database," *International Conference on Image Analysis and Processing*, pp. 970-977, 2005. http://iris.di.ubi.pt/

[9] W. S. Chen, Y. S. Chiang, C. M. Hsieh, S. W. Shih and Y. H. Chiu, "Biometric Verification with Hybrid Features of Palm-print and Hand Geometry based on Wavelet Transform and Statistical Moments," *IPPR Conference on Computer Vision, Graphics and Image Processing*, 2005.