

# Contents

Chapter 1	Fusion of Face and Palmprint for Personal Identification Based on Ordinal Feature	1
1.1	Introduction	1
1.2	Ordinal Features	4
1.2.1	Local Ordinal Features	5
1.2.2	Non-local Ordinal Features	6
1.3	Ordinal Feature Based Multimodal Biometric System	9
1.3.1	Face Recognition	10
1.3.2	Palmprint Recognition	11
1.4	Multimodal Biometric Fusion	13
1.4.1	Sum, Product, Max and Min rules	13
1.4.2	Linear Discriminant Analysis	15
1.5	Experiments	15
1.5.1	Data Description	15
1.5.2	Experimental Results and Evaluation	16
1.6	Conclusions	19



# Chapter 1

## Fusion of Face and Palmprint for Personal Identification Based on Ordinal Feature

Due to the limitations of universality and accuracy, the unimodal biometric systems are often unable to meet the high performance requirement imposed by large-scale authentication systems. In this chapter, we present a new multimodal biometric identification system in fusion of palmprint and face features such that the drawbacks of single biometric scheme are diminished and the identification performance is improved. Effective classifiers based on the powerful ordinal features are first constructed for palmprints and faces respectively. After that some strategies are employed for fusion of them on a middle-scale data set consists of 378 subjects and 20 pairs for each. Experimental results have demonstrated the effectiveness of the proposed system. <sup>1</sup>

### 1.1 INTRODUCTION

Biometrics, learning the physiological or behavioral characteristics of human being such as fingerprint, iris, face, palmprint, gait, and voice, has been acknowledged to provide advantages over the non-biometric methods such as password, PIN, and ID cards[1]. Its promising applications as well as the theoretical challenges have got its heat attraction from the last decade.

<sup>1</sup> The authors for this chapter are Rufeng Chu, Shengcai Liao, Yufei Han, Zhenan Sun, Stan Z. Li, Tieniu Tan. Center for Biometrics and Security Research & National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, 95 Zhongguancun East Road Beijing 100080, China

Most biometric systems deployed in real-world applications are unimodal relying on the evidence of a single source of biometrics information for authentication. However, due to the limitations of universality and accuracy, the unimodal biometric systems are often unable to meet the high performance requirement imposed by large-scale authentication systems[2]. For example, face image is influenced by illumination, pose and facial expression, and voiceprint is influenced by the environmental noise. Even though iris and fingerprint are reported to be highly accurate for authentication, however, for iris recognition the user must be cooperative making it difficult to acquire a suitable iris image, while for fingerprint about two percent of the population does not have a legible fingerprint and are impossible to be enrolled into a biometrics system[3]. Therefore, multimodal biometric systems integrating the evidence from multiple sources of information are developed to much more reliable performance. A number of studies showing the advantages of multimodal biometrics have appeared in the literature [4, 5, 6, 7, 8, 9, 10, 11, 12, 13].

In 1995, Brunelli and Falavigna[4] proposed a person identification system based on voice and face, and Kittler et al.[5] further evaluated such methodology using different fusion rules. As the fast development of biometrics techniques in the last decade, Hong and Jain [6] then proposed an identification system based on face and fingerprint and Wang et al. [7] proposed the fusion of face and iris. Recently, some new multimodal biometric systems using hand-based information for fusion are proposed. Kumar et al. [8] proposed a system combining the geometric features of the hand with palmprints. Kumar and Zhang [9] proposed an identification systems based on face and palmprint, where a feed-forward neural network is used to integrate individual matching scores and generate a combined decision score. Alternatively, Feng et al. [10] presented a face and palmprint multimodal biometric system by fusion of features extracted by PCA or ICA. Ribaric and Fratric [11] described a biometric identification system based on Eigenpalm and Eigenfinger features with fusion applied at the matching score level. Interestingly, some other approaches are also attractive, such as the fusion of ear and face proposed by Chang et al. [12] and the fusion of face, fingerprint, and hand geometry biometrics by Ross and Jain [13] etc.

Due to the high user acceptance, non-invasive and low-cost image acquisition device, face and palmprint multimodal biometrics have the significant advantages for personal identification. However, existing studies along this approach [9, 10] have employed holistic features for face representation, such as PCA and ICA, where only the experimental results on a small data set (less than 100 subjects) were reported. Note that PCA, ICA are typically sensitive to the global variation of face, such as illumination and inaccurate alignment. On the other hand, although

various palmprint representations have been proposed, such as Line features [16], Feature points [17], Fourier spectrum [18], Eigenpalms features [19], Sobel and morphological features [20], Texture energy [21], Wavelet signatures [22], Gabor phase [23], Fusion code [24], Competitive code [25] etc., what the proper representation for palmprint is should still be a question. Moreover how to model the palmprint pattern effectively and efficiently has so far not been well addressed.

For face recognition, in constant to the holistic features, local appearance features have been widely found to be useful and powerful for pattern recognition in which they are more stable to the global variation of pattern. In the last decade, some famous algorithms extracting local features, such as Local features analysis (LFA) [26], Gabor wavelet-based features [27, 28, 29] and local binary pattern (LBP) [30] etc., are developed. For palmprint recognition, much better performances are also reported by using local features[23, 24, 25].

It is believed that the human vision system uses a series of levels of representation, with increasing complexity. A recent study on local appearance or fragment (or local region) based object recognition [31] shows that features of intermediate complexity are optimal for basic visual task of classification, and mutual information for classification is maximized in a middle range of fragment size. Existing approaches suggest a trade-off between the complexity of features and the complexity of the classification scheme. Using fragment features is therefore advantageous [32] in that the number of features used for classification are reduced from richer information content of the individual features, and that a linear classifier may suffice when proper fragment features are selected, while for simple generic features the classifier has to use higher-order properties of their distributions.

In this regard, we consider a class of simple features: the ordinal relationship. Ordinal features are defined based on the qualitative relationship between two image regions and are robust against various intra-class variations [33, 34, 35]. For example, they invariant to monotonic transformations on images and is flexible enough to represent different local structures of different complexity. Sinha [34] shows that several ordinal measures on facial images, such as those between eye and forehead and between mouth and cheek, are invariant with different persons and imaging conditions, and thereby develops a ratio-template for face detection. Schneiderman [37] uses an ordinal representation for face detection.

Ordinal features also find other applications. Sinha exploited ordinal information of several attribute dimensions, such as intensity, color and shape, to construct a unique face signature of a scene [34]. Lipson *et al.* applied an ordinal technique to image database indexing [38]. Bhat and Nayar employed the relative intensity values in image windows for stereo correspondence [39]. By combining ordinal

measures and co-occurrence, Partio *et al.* obtained better texture retrieval results than traditional gray level co-occurrence matrices [40].

Along this line, in this chapter, we present a new multimodal biometric system for fusion of face and palmprint based on ordinal features. Ordinal features would be generated using ordinal filters for palmprints [41] and faces [42] respectively. Effective classifiers are then constructed for each unimodal biometric systems. Finally different strategies are employed for fusing palmprint and face classifiers on a relatively middle-scale data set. Though Thoresz [35] believed ordinal features may be only suited for simple detection and categorization but too weak for fine discrimination tasks, such as personal identification, however, our work presented here can break such view and show the power of ordinal feature. Experimental results have demonstrated the effectiveness of the proposed system.

The rest of this chapter is organized as follows. In Section 1.2, we introduce ordinal features. In Section 1.3, we present the system description and the details of the palmprint and face recognition methods. Section 1.4 describes the fusion methods of face and palmprint modalities. Experimental results are presented in Section 1.5.

## 1.2 ORDINAL FEATURES

Ordinal features come from a simple and straightforward concept that we often use. For example, we could easily rank or order the heights or weights of two persons, but it is hard to answer their precise differences. For computer vision, the absolute intensity information associated with an face can vary because it can changes under various illumination settings. However, ordinal relationships among neighborhood image pixels or regions present some stability with such changes and reflect the intrinsic natures of the object.

An ordinal feature encodes an ordinal relationship between two concept. Figure 1.1 gives an example in which the average intensities between regions A and B are compared to give the ordinal code of 1 or 0. Ordinal features are efficient to compute. Moreover, the information entropy of the measure is maximized because the ordinal code has nearly equal probability of being 1 or 0 for arbitrary patterns.

According to the spatial relationship between the image regions, Ordinal measure can be classified into two categories, local ordinal measure and non-local ordinal measure.



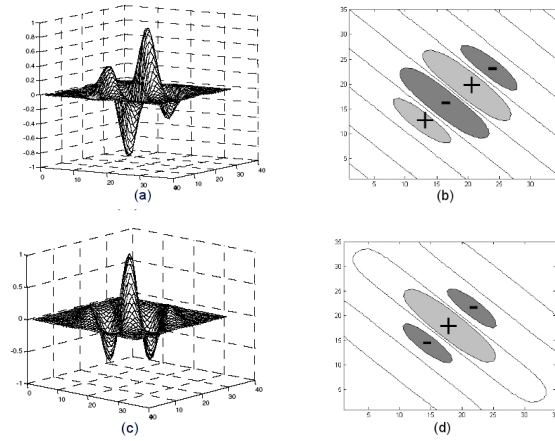
**Figure 1.1** Ordinal measure of relationship between two regions. An arrow points from the darker region to the brighter one. Left: Region A is darker than B, i.e.  $A < B$ . Right: Region A is brighter than B, i.e.  $A > B$ .

### 1.2.1 Local Ordinal Features

Local ordinal measure depict the comparison of adjacent image regions, which is well-suited to represent the object possessing a rich source of local sharp variations, such as iris and palmprint[36, 41]. A common practice to compare adjacent image regions is based on differential filters. After filtering, the local region covered by the operator is coded as 1 or 0 based on the sign of the filtering result.

In fact, many existing palmprint recognition methods [23, 24, 25] used this information implicitly [41]. For example, Gabor based encoding filters used in palm code [23] are essentially local ordinal operators (see Figure 1.2). For odd Gabor filtering of local palmprint region, the image regions covered by two excitatory lobes are compared with the image regions covered by two inhibitory lobes (Figure 1.2b). The filtered result is qualitatively encoded as 1 or 0 based on the sign of this inequality. Similarly, even Gabor generated palm code is mainly determined by the ordinal relationship between one excitatory lobe-covered region and two small inhibitory lobes-covered regions (Figure 1.2d). Because the sum of original even Gabor filters coefficients is not equal to 0, the average coefficient value is reduced from the filter to maximize the information content of the corresponding palm code.

In addition, ordinal relationship is not restricted to intensity measurement. As a byproduct of Gabor phase measure (i.e. ordinal intensity measure), the orientation energy or magnitude was also obtained by orthogonal Gabor filtering. Thus it is possible to combine ordinal intensity measures and ordinal energy measures together. In [24], the local energy along four different orientations were compared each other to obtain the maximum. Then the palmprint is represented using the Gabor filtered ordinal intensity measures whose basic lobes are along the maximum energy orientation. In [25], orientation of the dominant line segment is regarded as the palmprint feature. Even Gabor filter is used to filter the local image region along six different orientations, obtaining the corresponding contrast magnitudes.



**Figure 1.2** Odd and even Gabor filters used in [23]. (a) Odd Gabor filter. (b) Ordinal comparison of image regions using odd Gabor filter, + denotes excitatory lobe covered image region and - represents inhibitory lobe covered image region. (c) Even Gabor filter. (d) Ordinal comparison of image regions using even Gabor filter.

Based on the winner-take-all competitive rule, the index (ranging from 0 to 5) of the minimum contrast magnitude was represented by three bits, namely competitive code. Due to the success of these methods, we conclude that the ordinal measures are perhaps the most suitable representation for palmprint-based identification system.

### 1.2.2 Non-local Ordinal Features

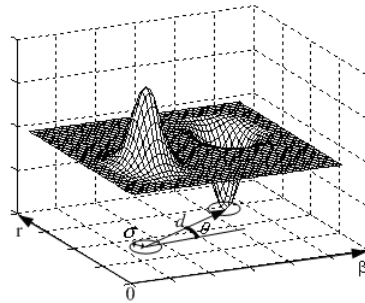
For face image, due to the similar facial shape of different persons, it is difficult to represent different faces by simple local ordinal comparison. Balas and Sinha [43] extend differential filters to “dissociated dipoles” for non-local comparison, which can compare small regions across large distances, shown in Figure 1.3. Like differential filters, a dissociated dipole also consists an excitatory and an inhibitory lobe, but the limitation on the relative position between the two lobes is removed. There are three parameters in dissociated dipoles:

- The scale parameter  $\sigma$ : On one hand, the noise suppression requires a coarse scale representation of the image structure. On the other hand, the discriminating power derives from fine details. The result of this trade-off is that an



intermediate scale should be carefully chosen or information at large scale and fine scale would be fused. For dipoles with a Gaussian filter, the standard deviation  $\sigma$  is an indicator of the scale.

- The inter-lobe distance  $d$ : This is defined as the distance between the centers of the two lobes. When  $d$  is equal to the size of the lobes in a dipole, the operator is essentially a local filter. So the differential filters can be seen as special case of dipoles.
- The orientation  $\theta$ : This is the angle between the line joining the centers of the two lobes and the horizontal line. It is in the range from 0 to  $2\pi$ .

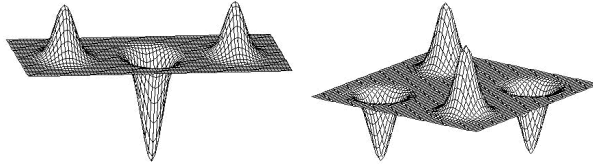


**Figure 1.3** Dissociated dipole operator.

We extend dissociated dipoles to dissociated multi-poles, as shown Figure 1.4. While a dipole tells us the orientation of a slope edge, a multi-pole can represent more complex image micro-structures. A multi-pole filter can be designed for a specific macro-structure, by using appropriate lobe shape configuration. This gives much flexibility for filter design.

To be effective for face recognition or image representation, there are three rules in development of dissociated multi-poles (DMPs):

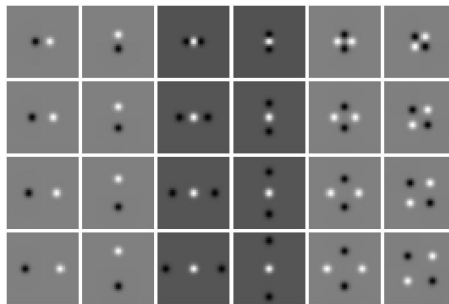
- Each lobe of a DMP should be a low-pass filter. On one hand, the intensity information within the region of the lobe should be statistically estimated; on the other hand, the image noise should be attenuated by low-pass filtering.
- To obtain the locality of the operator, the coefficients of each lobe should be arranged in such a way that the weight of a pixel is inverse proportional to



**Figure 1.4** Dissociated multi-pole: tri- and quad-pole filters.

its distance from the lobe center. Gaussian mask satisfies this; there are other choices as well.

- The sum of all lobes' coefficients should be zero, so that the ordinal code of a non-local comparison has equal probability being 1 or 0. Thus the entropy of a single ordinal code is maximized. In the examples shown in Figure 1.4, the sum of two excitatory lobes' weights is equal to the inhibitory lobes' total absolute weights.



**Figure 1.5** The 24 ordinal filters used in the experiments.

In our experiments, we design 24 disassociated multi-pole ordinal filters as shown in Figure 1.5. The filter sizes are all  $41 \times 41$  pixels. The Gaussian parameter is uniformly  $\sigma = \pi/2$ . The inter-pole distances are  $d = 8, 12, 16, 20$  for the 2-poles and 4-poles, and  $d = 4, 8, 12, 16$  for the 3-poles. For 2-poles and 3-poles, the directions are 0 and  $\pi/2$ ; for the 4-poles, the directions are 0 and  $\pi/4$ . Figure 1.6 shows the correspondingly filtered images of a face.

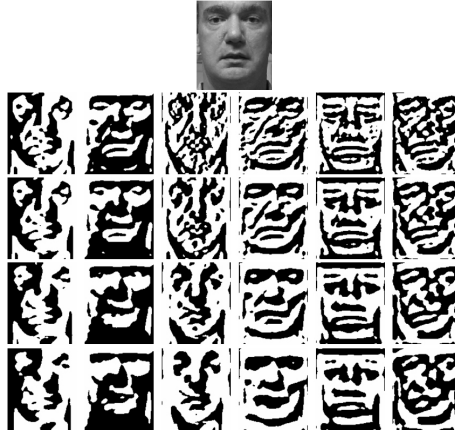


Figure 1.6 The filtered images of a face for the 24 ordinal filters.

### 1.3 ORDINAL FEATURE BASED MULTIMODAL BIOMETRIC SYSTEM

Figure 1.7 shows the block diagram of the proposed multimodal biometric system based on the fusion of face and palmprint at the matching score level. Firstly, effective face and palmprint ordinal features are extracted for matching. By comparing with the templates stored in the database, the matching scores of each classifier are generated. Then, the scores output from the two classifiers are combined using some fusion methods to give a unique matching score. Finally, a decision about whether to accept or reject a user is made.

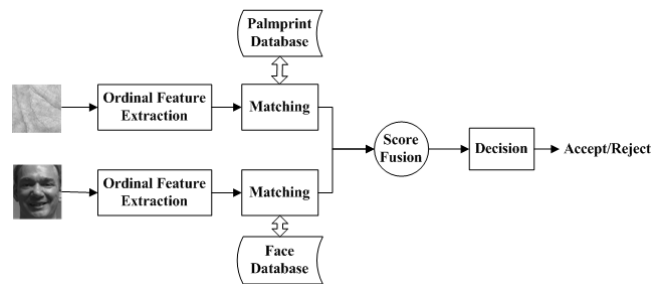


Figure 1.7 Block diagram of Face and palmprint multimodal biometric system.

### 1.3.1 Face Recognition

There are a large number of ordinal features generated by all ordinal types and all pixel locations, therefore the initial ordinal feature set is of high dimensionality. However, the intrinsic dimension of the face pattern may not be so high. A further processing is need to remove the redundancy and build effective classifier. This is done in this work by using the following AdaBoost algorithm [44]:

**Input:** Sequence of  $N$  weighted examples

$$\{(x_1, y_1, w_1), (x_2, y_2, w_2), \dots, (x_n, y_n, w_n)\};$$

Initial distribution  $P$  over the  $n$  examples;

Weak learning algorithm **WeakLearn**;

Integer  $T$  specifying number of iterations;

**Initialize**  $w_i^1 = P(i)$  for  $i = 1, \dots, n$ ;

For  $t = 1, \dots, T$ :

1. Set  $p_i^t = w_i^t / \sum_i w_i^t$ ;
2. Call **WeakLearn**, providing it with the distribution  $p$ ;  
get back hypothesis  $h_t(x_i) \in \{0, 1\}$  for each  $x_i$ ;
3. Calculate the error of  $h_t$  :  $\epsilon_t = \sum_{i=1}^N p_i^t |h_t(x_i) - y_i|$ ;
4. Set  $\beta_t = \frac{\epsilon_t}{(1-\epsilon_t)}$ ;
5. Set the new weights to  $w_i^{t+1} = \beta_t^{1-|h_t(x_i)-y_i|}$ ;

**Output** the hypothesis

$$H(x) = \begin{cases} 1 & \text{if } \sum_{t=1}^T \left(\log \frac{1}{\beta_t}\right) h_t(x) \geq \sum_{t=1}^T \left(\log \frac{1}{\beta_t}\right) \\ 0 & \text{otherwise} \end{cases}$$

AdaBoost iteratively learns a sequence of weak hypotheses  $h_t(x)$  and linearly combines them with the corresponding learned weights  $\log(1/\beta_t)$ . Given a data distribution  $p$ , AdaBoost assumes that a **WeakLearn** procedure is available for learning a sequence of most effective weak classifiers  $h_t(x)$ .

The simplest weak classifier can be constructed for each pixel and each filter type, called single bit weak classifier (*SBWC*). Consider the *SBWC* maybe unstable to image noise and alignment error, a more involved weak classifier can be designed based on a spatially local subwindow instead of a single bit. The advantage is that some statistic over a local subwindow can be more stable than that at a bit. In this scheme, the Hamming distance can be calculated between the ordinal values in

the two corresponding subwindows. The Hamming distance as a weak classifier can be used to make a weak decision for the classification. The use of subwindows gives one more dimension of freedom. A different size leads to a different weak classifier. In our experiment, 20 subwindow sizes are used:  $6 \times 6$ ,  $12 \times 12$ , ...,  $120 \times 120$  and the length of the side is incremented by 6.

### 1.3.2 Palmprint Recognition

Due to different illumination settings, stretching and misalignment, the acquired palmprint signal varies significantly. A robust palmprint image representation method is needed to ensure the high discriminant power. Although absolute intensity values of image regions are not reliable for recognition, their mutual ordinal relationships are stable against different variations.

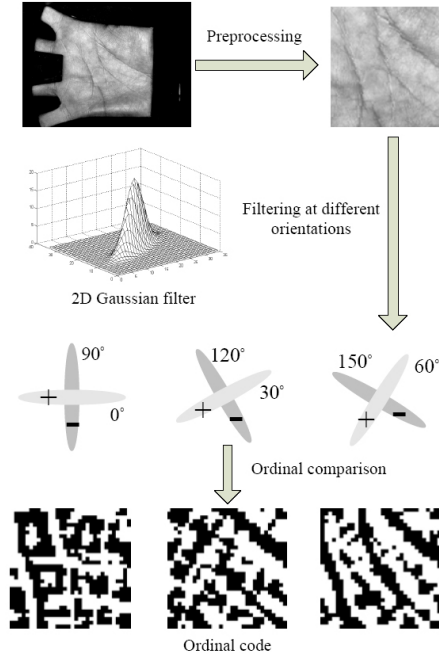
By analysis in section 1.2.1, a possible improvement could be made by choosing well-designed ordinal measures as the palmprint representation, into which the characteristics of palmprint pattern should be incorporated. We propose a novel palmprint representation, namely, *Orthogonal Line Ordinal Features (OLOF)*, as illustrated in Figure 1.8, where normalized subimage is referenced by finger gaps using an algorithm similar to Zhang *et al.*'s [23]. *OLOF* is so called because the two regions involved in ordinal comparison are elongated or line-like, and the two are geometrically orthogonal.

The ideas are motivated by the most stable and robust ordinal measures available in palmprint pattern, i.e. randomly distributed negative line segments versus their orthogonal regions. In low resolution palmprint images, the line patterns are mainly constituted by principal lines and wrinkles, whose intensity is much lower than their orthogonal regions. Of course detection of all line segments in palmprint is impossible in realtime applications. Nevertheless if we apply thousands of ordinal operators onto a palmprint image, most of them correspond to robust ordinal measures.

Here we use 2D Gaussian filter to obtain the weighted average intensity of a line-like region. Its expression is as follows:

$$f(x, y, \theta) = \exp\left\{-\frac{1}{2} \times \left[\left(\frac{x\cos\theta + y\sin\theta}{\delta_x}\right)^2 - \left(\frac{-x\sin\theta + y\cos\theta}{\delta_y}\right)^2\right]\right\} \quad (1.1)$$

where  $\theta$  denotes the orientation of 2D Gaussian filter,  $\delta_x$  denotes the filters horizontal scale and  $\delta_y$  denotes the filters vertical scale. We control the scale ratio  $\delta_x/\delta_y$  higher than 3 to make its shape like a line (see Figure 1.8).



**Figure 1.8** Orthogonal line ordinal features for palmprint recognition.

The orthogonal line ordinal filter, comparing two orthogonal line-like palmprint image regions, is specially designed as follows:

$$OF(\theta) = f(x, y, \theta) - f(x, y, \theta + \frac{\pi}{2}) \quad (1.2)$$

For each local region in normalized palmprint image, three ordinal filters,  $OF(0)$ ,  $OF(\pi/6)$ , and  $OF(\pi/3)$ , are performed on it to obtain three bit ordinal codes based on the sign of filtering results. Finally, three ordinal templates named as ordinal code are obtained as the feature of the input palmprint image (Figure 1.8). The matching metric is also based on Hamming distance.

## 1.4 MULTIMODAL BIOMETRIC FUSION

In the context of biometrics, there are mainly three levels of fusion methods for combining two (or more) biometric systems[13]: (a) fusion at the feature extraction level, where the feature vectors of multiple biometric modalities are concatenated to create a new feature vector to represent the individual, (b) fusion at the matching score level, where the matching scores of multiple classifiers are combined to generate a single scalar score, (c) fusion at the decision level, where each matcher outputs its own Boolean result and the fusion process fuses them together by a combination algorithm such as AND, OR, etc.

In our multimodal biometric system, the fusion is performed at the matching score level. Kittler et al. [45] developed a common theoretical framework for combining classifiers and showed that many existing schemes could be considered as special cases of compound classification where all the pattern representations are used jointly to make a decision. The fusion procedure can be formulated as follows[45]. Given  $s_1$  and  $s_2$  (matching score of each classifier), a tester  $T$  is assigned to one of the two possible classes  $\omega_1$ (genuine) and  $\omega_2$ (imposter), i.e.

$$\begin{aligned} & \text{assign } T \rightarrow \omega_j, \text{ if} \\ P(\omega_j|s_1, s_2) &= \max_k P(\omega_k|s_1, s_2), \quad j = 1, 2 \end{aligned} \quad (1.3)$$

where  $P(\omega_j|s_1, s_2)$  are the posteriori of classes  $\omega_j$ .

There are two approaches for consolidating the scores obtained from different matchers. One is to formulate it as a combination problem and use simple fusion rules, such as sum, product, max and min rule[45], to combine the two matching scores and compare the result to a threshold. The other is to formulate it as a classification problem and treat the matching scores of different biometrics as a feature vector and use linear discriminant analysis (LDA) to classify the vector as being genuine or an impostor.

### 1.4.1 Sum, Product, Max and Min rules

In the fusion approach, the individual matching scores are combined to generate a single scalar score, which is then used to make the final decision. Let  $S_1$  and  $S_2$  be the matching scores generated by face and palmprint classifier respectively. To ensure a meaningful combination of the scores from different modalities, the scores must be first transformed to a common domain prior to combining them. This is known as score normalization. In our experiments, the following normalization

techniques are used: simple Min-Max, Z-Score and Hyperbolic tangent (Tanh) [46]. Note that all the following score normalization methods are only for the combination approaches. The definitions are as follows:

- Min-max normalization is used for normalizing each of the matching scores to a scale of 0 to 1,

$$S'_i = \frac{S_i - \min(S_i)}{\max(S_i) - \min(S_i)}, \quad i = 1, 2 \quad (1.4)$$

where  $\min(S_i)$  and  $\max(S_i)$  denote the overall minimum and maximum value of  $S_i$ .

- Z-Score normalization method transforms the scores to a distribution with mean of 0 and standard deviation of 1.

$$S'_i = \frac{S_i - \mu_i}{\sigma_i}, \quad i = 1, 2 \quad (1.5)$$

where  $\mu_i$  and  $\sigma_i$  denote the arithmetic mean and standard deviation, respectively:

- Tanh normalization method is a stable statistical techniques [47]. It maps the raw scores to the (0, 1) range:

$$S'_i = \frac{1}{2} \left[ \tanh\left(0.01 \frac{S_i - \mu_i}{\sigma_i}\right) + 1 \right], \quad i = 1, 2 \quad (1.6)$$

After the score normalization, the final fused matching scores can be obtained using the following strategies:

- Sum rule:

$$S = (S'_1 + S'_2)/2 \quad (1.7)$$

- Product rule:

$$S = S'_1 \cdot S'_2 \quad (1.8)$$

- Max rule:

$$S = \max\{S'_1, S'_2\} \quad (1.9)$$

- Min rule:

$$S = \min\{S'_1, S'_2\} \quad (1.10)$$



### 1.4.2 Linear Discriminant Analysis

Linear discriminant analysis(LDA) helps transform the two-dimensional score vectors  $\vec{s} = (s_1, s_2)$  into a new subspace that maximizes the between-class distance while minimized the within-class distance. Consider a set of feature vectors  $\vec{s}$  for each sample of an object with known class labels  $\omega_j (j = 1, 2)$ . The classification problem is then to find a good predictor of the class  $\omega_j$  for any any observation  $\vec{s}$ .

LDA is known to be Bayes optimal if the probability density functions  $p(\vec{s}|\omega_1)$  and  $p(\vec{s}|\omega_2)$  are both normally distributed with equal covariance matrix. Using LDA, one can measure the probability  $P(\omega_j|\vec{s})$  by calculating the distance between the observation and the class prototype in a discriminant subspace induced by the following projection [48].

$$\vec{W} = \Sigma^{-1}(m_1 - m_2) \quad (1.11)$$

$$\vec{m}_i = \frac{1}{n_i} \sum_{\vec{s} \in D_i} \vec{s} \quad (1.12)$$

$$\Sigma = \sum_{i=1,2} \sum_{\vec{s} \in D_i} (\vec{s} - \vec{m}_i)(\vec{s} - \vec{m}_i)^T \quad (1.13)$$

where  $\Sigma$  and  $\vec{m}_i$  denote the covariance and mean of the input feature vector respectively, and  $D_1$  and  $D_2$  denote the set of samples from the genuine and impostor classes respectively.

## 1.5 EXPERIMENTS

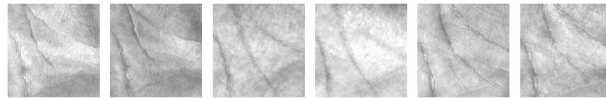
### 1.5.1 Data Description

We evaluate the proposed multimodal system on a database of 7560 pairs of images from 378 subjects. The face images are collected from FRGC 2.0 face database[50] which contains 466 subjects and about 36818 still images. Of the 466 distinct subjects, the number of images for each subject varies from 4 to 88, coming from several sessions for each subject. We selected 7560 images from 378 different subjects with 20 samples for each. All these face images are normalized to  $142 \times 120$  and preprocessed using the method[51]. Face images in this subset mainly subject to illumination, poses and facial expressions. Examples of typical face images from FRGC 2.0 database are shown in Figure 1.9.



**Figure 1.9** Sample face images from FRGC2.0 database

The palmprint images are collected from PolyU Palmprint Database[49] which contains 7560 images corresponding to 378 different palms (hence 378 classes), where 20 samples from each of these palms were collected in two sessions, i.e., 10 samples were captured in the each session respectively. The average interval between those two collections was two months. After preprocessing, the input palmprint image is normalized to  $128 \times 128$ . The variations in the palmprint images were mainly due to illumination, stretching and misalignment. Some palmprint images are shown in Figure 1.10.



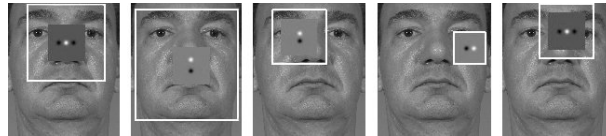
**Figure 1.10** Sample face images from PolyU Palmprint Database

Though face images in FRGC 2.0 database do not correspond to the the palmprint images, Ross *et al.* [13] showed that biometric modalities were mutually independent. This allows us to randomly pair face and palmprint to obtain a multimodal image set for each of the 378 subjects. Each pair contains 20 face images and 20 palmprint images. In our experiments, we divide the images into two partitions. The first 3780 pairs of images including 187 subjects and 20 pairs of images for each, are used for training, and the rest for testing.

### 1.5.2 Experimental Results and Evaluation

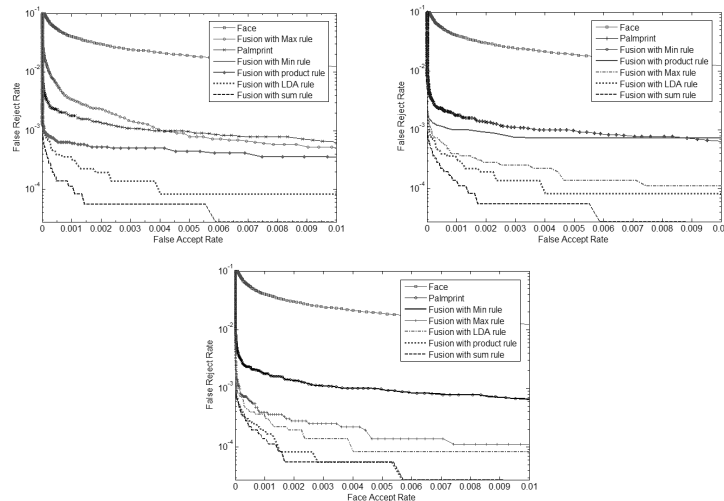
As for palmprint matching, a training procedure is actually not required. Each of palmprint image was directly used to obtain a characteristic, i.e., ordinal feature vector of size being 384 bytes. While for face matching, each of the image was filtered by 24 ordinal filters. Discriminant analysis is then applied to screen out less effective features, and local subwindow of ordinal features are used to construct weak classifier based on Hamming distance for AdaBoost Learning. Finally a strong

classifiers constructed by 2014 weak classifiers is obtained. For illustration, the first 5 learned weak classifiers are shown in Figure 1.11.



**Figure 1.11** The first 5 features and associated subwindow sizes selected by AdaBoost learning.

In both training and testing set, there are totally 35,910 intra-class (genuine) samples and 7,106,400 extra-class (impostor) samples for each. The genuine and impostor matching scores from the training set were used to learn the LDA classifier. Then we test the performance of face, palmprint and different type of fusion classifiers on the testing set. Figure 1.12 shows the ROC curves of different classifiers derived from the scores for the intra- and extra-class pairs.



**Figure 1.12** ROC curves for comparing fusion methods with different normalization methods: (a) Min-max; (b) Z-Score; (c) Tanh.

The ROC curves suggest that the multimodal biometric offers substantial performance gain. Only the max rule performs slightly worse than a single palmprint

classifier when using min-max score normalization method. In particular, it is found that fusion with sum rule method has the highest accuracy for all score normalization techniques. The performance of LDA classifier is similar to that of the sum rule based one. The performance of product rule is not stable if different score normalization methods are utilized.

For further comparison with different fusion methods, two measurements are employed. One is the equal error rate (EER) described by a point in ROC when false accept rate is equal to false reject rate and another is the  $d'$  (d-prime)[52], which are shown in Table 1.1. Where  $d'$  is a statistical measure of how well a biometric system can discriminate between different individuals. The definition is

$$d' = \frac{|m_1 - m_2|}{\sqrt{(\delta_1^2 + \delta_2^2)/2}} \quad (1.14)$$

where  $m_1$  and  $\delta_1$  denote the mean and variance of intra-class feature vector respectively while  $m_2$  and  $\delta_2$  denote the mean and variance of extra-class feature vector respectively. The larger the  $d'$  value is, the better a biometric system performs at discriminating between individuals.

**Table 1.1**  
Comparison of accuracy measures for different classifiers

Algorithm	Score normalization method	EER (%)	$d'$
Face	-	1.160	4.4033
Palmprint	-	0.160	6.0813
Fusion with sum rule	Min-max	<b>0.028</b>	<b>7.3916</b>
	Z-Score	0.032	7.3645
	Tanh	0.032	7.3723
Fusion with product rule	Min-max	0.067	5.7714
	Z-Score	0.091	4.0360
	Tanh	0.032	7.2822
Fusion with max rule	Min-max	0.220	5.4496
	Z-Score	0.056	6.2805
	Tanh	0.056	6.2903
Fusion with min rule	Min-max	0.150	6.7518
	Z-Score	0.160	5.5558
	Tanh	0.150	5.5598
Fusion with LDA	-	0.044	7.1244

Experimental results in Table 1.1 indicate that the sum rule based method achieves better performance. Similar conclusion can be drawn for LDA. For other

fusion rules, we see that the product rule is good for improvement of EER which is significant but the discriminating index drops a little when min-max or Z-Score normalization is employed, while for the max rule all indices are better than that of a single face or palmprint classifier when Z-Score or Tanh normalization is utilized.

## 1.6 CONCLUSIONS

We have presented a multimodal biometric identification system in fusion of palmprint and face, which takes advantages of ordinal measure for palmprint and face representations. Our work has shown that well-designed ordinal features can be powerful enough for complex tasks such as personal identification, and this has broken the existing view that ordinal features are believed to be only suited for face detection and too weak for fine discrimination tasks [35].

In particular, we have investigated the multimodal fusion methods at score-level, and examined the multimodal biometric system on a middle-scale population (378 subjects, 20 pairs each). The experimental results show that the proposed system significantly improves the performance of identification system. Furthermore, our score-level fusion experiments show that sum rule and LDA generally perform better than min rule and max rule. Thanks to the high accuracy and low computation cost, the best combination is a simple sum rule with min-max score normalization.

In the future, we will further investigate the performance of the proposed system on a larger database. In addition, we will explore the feature-level fusion based on ordinal features and statistical learning and compare the performances of different level fusion methods.

## References

- [1] Jain, A.K., R.M. Bolle, and Eds S. Pankanti. *Biometrics: Personal Identification in Networked Society*. Kluwer, 1999.
- [2] Ross, A., and A. K. Jain. "Multimodal Biometrics: An Overview". In *Proc. of 12th European Signal Processing Conference (EUSIPCO)*, pages 1221–1224, 2004.
- [3] NIST Report. *Summary of NIST Standards for Biometric Accuracy, Tamper Resistance, and Interoperability*. NIST Report to the United States Congress, Nov. 2002.
- [4] Brunelli R., and D. Falavigna. "Person identification using multiple cues". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:955–966, 1995.
- [5] Kittler, J., et al. "Combining evidence in personal identity verification systems". *Pattern Recognition Letters*, 18:845–852, 1997.

- [6] Hong L., and A. K. Jain. "Integrating faces and fingerprints for personal identification". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:1295–1307, 1998.
- [7] Wang, Y., T. Tan, and A. K. Jain. "Combining face and iris biometrics for identity verification". In *Proc. of Int'l Conf on Audio- and Video-based Person Authentication*, pages 805–813, 2003.
- [8] Kumarand, A., et al. "Personal verification using palmprint and hand geometry biometric". In *Proc. of Int'l Conf on Audio- and Video-based Person Authentication*, pages 668–678, 2003.
- [9] Kumar, A., and D. Zhang. "Integrating palmprint with face for user authentication". In *Proc. Multi Modal User Authentication Workshop*, pages 107–112, 2003.
- [10] Feng, G., et al. "When Faces Are Combined with Palmprints: A Novel Biometric Fusion Strategy". In *Proceedings of International Conference on Biometric Authentication*, pages 701–707, 2004.
- [11] Ribaric, S., and I. Fratric. "A Biometric identification system based on eigenpalm and eigenfinger features". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27:1698–1709, 2005.
- [12] Chang, K., et al. "Comparison and combination of ear and face images in appearance based biometrics". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:1160–1165, 2003.
- [13] Ross, A., A. K. Jain, and J. Z. Qian. "Information fusion in biometrics". *Pattern Recognition Letters*, 24:2115–2125, 2003.
- [14] Turk, M. A., and A. P. Pentland. "Eigenfaces for recognition". *Journal of Cognitive Neuroscience*, 3(1):71–86, March 1991.
- [15] Bartlett, M. S., H. Martin Lades, and T. J. Sejnowski. "Independent component representations for face recognition". *Proceedings of the SPIE, Conference on Human Vision and Electronic Imaging III*, 3299:528–539, 1998.
- [16] Zhang, D., and W. Shu. "Two novel characteristics in palmprint verification: Datum point invariance and line feature matching". *Pattern Recognition*, 32:691–702, 1999.
- [17] Duta, N., A. K. Jain, and K. V. Mardia. "Matching of palmprint". *Pattern Recognition Letters*, 23:477–485, 2001.
- [18] Li, W., D. Zhang, and Z. Xu. "Palmprint identification by fourier transform". *International Journal of Pattern Recognition and Artificial Intelligence*, 16:417–432, 2002.
- [19] Lu, G., D. Zhang, and K. Wang. "Palmprint recognition using eigenpalms features". *Pattern Recognition Letters*, 24:1463–1467, 2003.
- [20] Han, C. C., et al. "Personal authentication using palmprint features". *Pattern Recognition*, 36:371–381, 2003.
- [21] You, J., et al. "On hierarchical palmprint coding with multi-features for personal identification in large databases". *IEEE Transactions on Circuits and Systems for Video Technology*, 14:234–243, 2004.
- [22] Zhang, L., and D. Zhang. "Characterization of palmprints by wavelet signatures via directional context modeling". *IEEE Transactions on Systems, Man and Cybernetics, Part B*, 34:1335–1347, 2004.
- [23] Zhang, D., et al. "On-line palmprint identification". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25:1041–1050, 2003.

- [24] Kong,W.K., and D. Zhang. "Feature-Level Fusion for Effective Palmprint Authentication". In *Proceedings of International Conference on Biometric Authentication*, pages 761–767, 2004.
- [25] Kong,W.K., and D. Zhang. "Competitive coding scheme for palmprint verification". In *Proceedings of International Conference Pattern Recognition*, volume 1, pages 520–523, 2004.
- [26] Penev,P., and J. Atick. "Local feature analysis: A general statistical theory for object representation". *Neural Systems*, 7(3):477–500, 1996.
- [27] Lades,M., et al. "Distortion invariant object recognition in the dynamic link architecture". *IEEE Transactions on Computers*, 42:300–311, 1993.
- [28] Wiskott,L., et al. "Face recognition by elastic bunch graph matching". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):775–779, 1997.
- [29] Liu,C., and H. Wechsler. "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition". *IEEE Transactions on Image Processing*, 11(4):467–476, 2002.
- [30] Ahonen,T., A. Hadid, and M.Pietikainen. "Face recognition with local binary patterns". In *Proceedings of the European Conference on Computer Vision*, pages 469–481, Prague, Czech, 2004.
- [31] Ullman,S., M. Vidal-Naquet, and E. Sali. "Visual features of intermediate complexity and their use in classification". *Nature Neuroscience*, 5(7), 2002.
- [32] Naquet,M.V., and S. Ullman. "Object recognition with informative features and linear classification". In *Proceedings of IEEE International Conference on Computer Vision*, Nice, France, 2003.
- [33] Sadr,J., et al. "Toward the fidelity of local ordinal encoding". In *Proceedings of the Fifteenth Annual Conference on Neural Information Processing Systems*, Vancouver, British Columbia, Canada, December 3-8 2001.
- [34] Sinha,P., "Toward qualitative representations for recognition". In *Proceedings of the Second International Workshop on Biologically Motivated Computer Vision*, pages 249–262, Tubingen, Germany, November 22-24 2002.
- [35] Thoresz.K. J, "On qualitative representations for recognition". Master's thesis, MIT, July 2002.
- [36] Sun,Z., T. Tan, and Y. Wang. "Toward Robust Encoding of Local Ordinal Measures: A General Framework of Iris Recognition ". In *Proceedings of International Workshop on Biometric Authentication*, pages 270-282, Prague, Czech Republic, May 2004.
- [37] Schneiderman. H., "Toward feature-centric evaluation for efficient cascaded object detection". In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1007–1013, Washington, DC, USA, June 27 - July 2 2004.
- [38] Lipson,P., E. Grimson, and P. Sinha. "Toward configuration based scene classification and image indexing". In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1007–1013, San Juan, Puerto Rico, June 17-19 1997.
- [39] Bhat,D., and S. Nayar. "On ordinal measures for image correspondence ". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(4):415–423, April 1998.
- [40] Partio,M., B. Cramariuc, and M. Gabbouj. "Toward texture similarity evaluation using ordinal co-occurrence". In *Proceedings of IEEE International Conference on Image Processing*, pages 1537–1540, Singapore, October 24-27 2004.

- [41] Sun,Z., T. Tan, Y. Wang, and S. Z. Li. "Ordinal palmprint representation for personal identification". In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 279–284, 20-25 June 2005.
- [42] Liao,S., et al. "Face recognition using ordinal features". In *Proceedings of IAPR International Conference on Biometric*, pages 40–46, Hong Kong, Jan 2006.
- [43] Balas,B., and P. Sinha. "Toward dissociated dipoles: Image representation via non-local comparisons". CBCL Paper #229/AI Memo #2003-018, MIT Computer Science and Artificial Intelligence Laboratory, Cambridge, MA, USA, August 2003.
- [44] Freund,Y., and R.E. Schapire. "A decision-theoretic generalization of on-line learning and an application to boosting". *Journal of Computer and System Sciences*, 55(1):119–139, August 1997.
- [45] Kittler,J., et al. "On combining classifiers". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(3):226–239, January 1998.
- [46] Samoska.N., "Evaluation and performance prediction of multimodal biometric systems". Master's thesis, West Virginia University, 2006.
- [47] Huber.P., *Robust Statistics*. Wiley, 1981.
- [48] Duda,R. O., P. E. Hart, and D. G. Stork. *Pattern Classification, Second Edition*. Wiley, 2000.
- [49] PolyU Palmprint Database. <http://www.comp.polyu.edu.hk/~biometrics/>.
- [50] Phillips,P.J., et al. J. Marques, J. Min, and W. Worek. "Overview of the face recognition grand challenge". In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.
- [51] Gross R., and V. Brajovic. "An image preprocessing algorithm for illumination invariant face recognition". In *Proc. 4th International Conference on Audio- and Video-Based Biometric Person Authentication*, pages 10–18, Guildford, UK, June 9 - 11 2003.
- [52] Daugman J., and G. Williams. "A Proposed Standard for Biometric Decidability". In *Proc. CardTech/SecureTech Conference*, pages 223–234, 1996.



# Index

Face recognition, 10

Linear discriminant analysis, 15

Max rule, 14

Min rule, 14

Multimodal biometric fusion, 13

    decision level, 13

    feature extraction level, 13

    matching score level, 13

Ordinal features, 4

    Local ordinal measure, 5

    Non-local ordinal measure, 6

    Orthogonal line ordinal features, 11

Palmprint recognition, 11

Product rule, 14

Score normalization, 14

Sum rule, 14