A Face and Palmprint Recognition Approach Based on Discriminant DCT Feature Extraction

Xiao-Yuan Jing and David Zhang

Abstract—In the field of image processing and recognition, discrete cosine transform (DCT) and linear discrimination are two widely used techniques. Based on them, we present a new face and palmprint recognition approach in this paper. It first uses a two-dimensional separability judgment to select the DCT frequency bands with favorable linear separability. Then from the selected bands, it extracts the linear discriminative features by an improved Fisherface method and performs the classification by the nearest neighbor classifier. We detailedly analyze theoretical advantages of our approach in feature extraction. The experiments on face databases and palmprint database demonstrate that compared to the state-of-the-art linear discrimination methods, our approach obtains better classification performance. It can significantly improve the recognition rates for face and palmprint data and effectively reduce the dimension of feature space.

Index Terms—Discrete cosine transform (DCT), DCT frequency band selection, improved Fisherface method, linear discrimination technique, two-dimensional (2-D) separability judgment.

I. INTRODUCTION

FREQUENCY-DOMAIN analysis is a commonly used image processing and recognition technique. During the past years, some work has been done to extract the frequency-domain features for image recognition. Li *et al.* extract Fourier range and angle features to identify the palmprint image [1]. Lai *et al.* use holistic Fourier invariant features to recognize the facial image [2]. Another spectral feature generated from singular value decomposition (SVD) is used by some researchers [3]. However, Tian *et al.* indicate that this feature does not contain adequate information for face recognition [4]. In [5], Hafed and Levine extract discrete cosine transform (DCT) feature for face recognition. They point out that DCT obtains the near-optimal performance of Karhunen–Loeve (KL) transform in facial information compression. And the performance of DCT is superior to those of discrete Fourier transform (FT) and other conventional transforms. By manually selecting the frequency bands of DCT, their recognition method achieves similar recognition effect as the Eigenface method [6]

which is based on KL transform. Nevertheless, their method cannot provide a rational band selection rule or strategy. And it cannot outperform the classical Eigenface method.

Image data are always high-dimensional and require considerable computing time for classification. Linear discrimination technique is thus, important in extracting effective discriminative features and reducing dimensionality of image. This technique usually needs much less computational cost than nonlinear recognition techniques like neural network. So far many linear discrimination methods have been proposed for use in image recognition. Two of the most well known are the Eigenface and Fisherface methods [7].

The Eigenface method uses the total covariance (or scatter) matrix S_t, as the production matrix to perform the KL transform. It cannot, however, make full use of pattern separability information like the Fisher criterion, and its recognition effect is not ideal when the size of the sample set is large [8]. Based on the Eigenface method and the Fisher criterion [9], the Fisherface method extracts the information that discriminates between the classes of a sample set. Nevertheless, Martinez and Kak demonstrate that when the training data set is small, the Eigenface method outperforms the Fisherface method [8]. Should the latter be outperformed by the former? This provoked a variety of explanations. Liu and Wechsler think that it might have been because the Fisherface method uses all the principal components, but the components with the small eigenvalues correspond to high-frequency components and usually encode noise [10], leading to recognition results that are less than ideal. In line with this theory, they present two enhanced Fisher linear discrimination (FLD) models (EFM) [10] and an enhanced Fisher classifier [11] for face recognition. Their explanation lacks sufficient theoretical demonstration, however, and EFM does not provide an automatic strategy for selecting the components.

In recent years, researchers have proposed some new ideas to further develop them theoretically. Chen *et al.* prove that the null space of the within-class scatter matrix S_w contains the most discriminative information when a small sample size problem takes place [12]. Their method is also inadequate, however, as it does not use any of the information outside the null space. In [13], Yu and Yang propose a direct linear discrimination analysis (DLDA) approach to solve this problem. It simultaneously diagonalizes both the between-class scatter matrix S_b (or S_t) and the within-class scatter matrix S_w. The shortcoming of DLDA is that it is rather difficult to demonstrate how to select the eigenvectors of D_w, which is necessary in theory. Chien and Wu present a discriminant waveletface method for face recognition [14], which extracts linear discriminative feature from a

specific level of low-frequency subimages of wavelet decomposition and uses nearest feature classifier to perform the classification. Nevertheless, it cannot theoretically determine which level of subimage is most appropriate for extracting the linear discriminative features. Moreover, it cannot make use of the discriminant information from the edge subimages of wavelet decomposition.

In addition, some extended discrimination methods are proposed. Zhang *et al.* present a dual Eigenspace method for face recognition [15]. In [16] and [17], several new discrimination principles based on the Fisher criterion were proposed. Yang use kernel principal component analysis (PCA) for facial feature extraction and recognition [18], while Bartlett *et al.* apply the independent component analysis (ICA) in face recognition [19]. However, Yang shows that both ICA and kernel PCA need much more computing time than PCA. In addition, when the Euclidean distance is used, there is no significant difference in the classification performance of PCA and ICA [18]. Jing *et al.* put forward a classifier combination method for face recognition [24]. In this paper, we do not analyze and compare these extended discrimination methods, and we confine ourselves to a comparison of major linear discrimination methods including the Eigenface method, the Fisherface method, DLDA and discriminant waveletface.

To enhance the image classification information and improve the recognition effect, we propose a new image recognition approach, which combines DCT with the linear discrimination technique. It first uses a two-dimensional (2-D) separability judgment that can facilitate the selection of useful DCT frequency bands for image recognition, because not all the bands are useful in classification. It will then extract the linear discriminative features by an improved Fisherface method and perform the classification by the nearest neighbor classifier. We will perform the detailed analysis of the theoretical advantages of our approach. The rest of this paper is organized as follows. In Section II, we provide the description of our approach. Section III shows its theoretical analysis. Next, the experimental results on different image data and some conclusions are given in Sections IV and V, respectively.

II. APPROACH DESCRIPTION

In this section, we present a 2-D separability judgment and introduce the whole recognition procedure of our approach.

A. Select DCT Frequency Bands by Using a 2-D Separability Judgment

Suppose that image training and test sample sets are X_1 and X_2, respectively, each gray image matrix is sized $M \times N$ and expressed by $f(x, y)$, where $1 \le x \le M$, $1 \le y \le N$ and $M \ge N$. Assume there are c known pattern classes $(w_1, w_2, \cdots, w_c)$ in X, where $P_i(i = 1, 2, \cdots, c)$ denotes the a priori probability of class w_i. Perform a 2-D DCT on each image [5] by

$$F(u, v) = \frac{1}{\sqrt{MN}} \alpha(u)\alpha(v) \sum_{x=1}^{M} \sum_{y=1}^{N} f(x, y)$$
$$\times \cos\left[\frac{(2x+1)u\pi}{2M}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \quad (1)$$

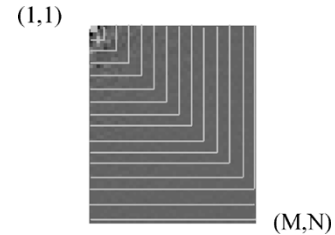Fig. 1. Demo of a facial image and its DCT transformed image.

Fig. 2. Illustration of expression ways of DCT frequency bands.

where $F(u, v)$ is sized $M \times N$, and $\alpha(\bullet)$ is defined as follows:

$$\alpha(w) = \begin{cases} \frac{1}{\sqrt{2}}, & w = 1 \\ 1, & \text{otherwise.} \end{cases} \quad (2)$$

Fig. 1(a) and (b) represents a facial image and its transformed image. From Fig. 1(b), most information or energy of image is concentrated in the left-up corner, that is, in the low frequency bands. Here, we provide a 2-D expression for different bands of the transformed image. A half square ring $\text{Ring}(k)$ is used to represent the kth frequency band. Different DCT frequency bands with the above expression ways are illustrated in Fig. 2. When $k \le N$, the three vertexes of $\text{Ring}(k)$ are $(1, k)$, $(k, 1)$ and (k, k), respectively. When $N < k \le M$, $\text{Ring}(k)$ is represented by only one side whose two vertexes are $(k, 1)$ and (k, N), respectively. So, the kth frequency band denotes

$$F(u, v) \in \text{Ring}(k), \quad 1 \le k \le M. \quad (3)$$

If we select the kth frequency band, then we keep the original values of $F(u, v)$, otherwise set the values of $F(u, v)$ to change to zero. Which principle should we follow to select the appropriate bands? Here, we propose a 2-D separability judgment to evaluate the separability of the frequency bands and select them.

1) Use the kth frequency band

$$F(u, v) = \begin{cases} \text{Original values}, & \text{if } F(u, v) \in \text{Ring}(k) \\ 0, & \text{if } F(u, v) \notin \text{Ring}(k). \end{cases} \quad (4)$$

Thus, for the images in X_1, we obtain the corresponding band-pass filtered images $F(u, v)$, which construct a new 2-D sample set Y_k. Obviously, Y_k and X_1 represent the same numbers of classes, the number of samples and the priori probabilities. Assuming that $A_i(i = 1, 2, \cdots, c)$
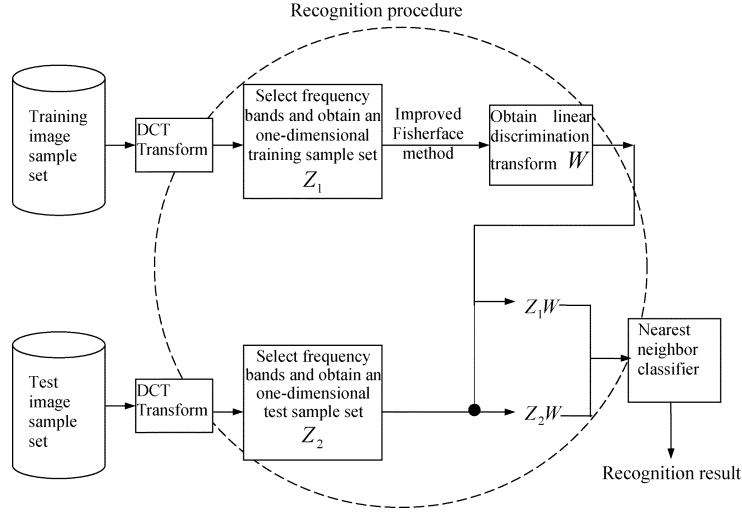
Fig. 3. Illustration of the image recognition procedure of our approach.

denotes a mathematic expected value of class w_i in Y_k and A denotes the total expected value of Y_k

$$A_i = E\left(Y_k(w_i)\right) \text{ and } A = E(Y_k). \tag{5}$$

Here, A_i and A are 2-D matrices whose dimensions are corresponding to u and v in (3). For Y_k, the between-class scatter matrix S_b, the within-class scatter matrix S_w and the total scatter matrix S_t are defined as

$$S_b = \sum_{i=1}^{c} P_i \left[(A_i - A)(A_i - A)^T\right] \tag{6}$$

$$S_w = \sum_{i=1}^{c} P_i E \left[(Y_k - A_i)(Y_k - A_i)^T\right] \tag{7}$$

$$S_t = E \left[(Y_k - A)(Y_k - A)^T\right] = S_b + S_w. \tag{8}$$

2) We evaluate the separability of Y_k, $J(Y_k)$, using the following judgment:

$$J(Y_k) = \frac{tr(S_b)}{tr(S_w)} \tag{9}$$

where $tr(\ )$ represents the trace of the matrix. For all the frequency bands $(1 \leq k \leq M)$, we select the bands by

$$J(Y_k) > T_1. \tag{10}$$

When $T_1 = 1$, that is $tr(S_b) > tr(S_w)$. There is more between-class separable information than within-class scatter information for Y_k according to the Fisher criterion. In other words, the corresponding selected frequency band has good linear separability. Hence, the theoretical value of T_1 should be 1.0. However, its experimental value might not be completely consistent with its theoretical value. In the experiments, we tune the experimental value of T_1 according to different data. The data with fewer samples often has fewer frequency bands whose separability values are more than 1. So, T_1 is set at less than 1.0 in order to use the bands with comparatively higher separability values as much as possible. The data with more samples often has more frequency bands

whose separability values are more than 1. So, T_1 is set more than 1.0 in order to select the most effective bands from many candidates.

We obtain a 2-D training sample set Y with all the selected bands. Note that Y_k is corresponding to only one selected band, i.e. the kth frequency band, but Y is corresponding to all selected bands. It should have favorable total separability value $J(Y)$, which can be similarly computed by (9). The experiments will show that $J(Y)$ obtained after band selection is greater than that obtained before selection. Notice that if we only use one frequency band with the maximum of $J(Y_k)$, it is difficult to guarantee that the selected band has good generalization capability in classification, because the number of training image samples is always very limited. Therefore, for image recognition, a range of frequency bands should be selected.

B. Recognition Procedure

Fig. 3 illustrates the whole procedure of image recognition. We first select the appropriate frequency bands for the training sample set, then an improved Fisherface method is proposed to extract the image discrimination features and the nearest neighbor classifier is applied to the feature classification:

Step 1) Use the measure introduced in Section II-A to select the appropriate frequency bands. If the kth frequency band is selected, then all $F(u, v)$ values belonging to this band are kept and represented by a feature vector. Then, we link all the feature vectors to a vector. In other words, each sample is represented by a feature vector. Thus we obtain a one–dimensional (1-D) training sample set Z_1 corresponding to X_1 and Y. We can also acquire a 1-D test sample set Z_2 corresponding to X_2. For Z_1, compute its S_b, S_w and S_t.

Step 2) Perform the following improvements on the original Fisherface method:

a) Calculate the discriminant transform W_{opt}:

$$W_{\text{opt}} = W_{\text{pca}} W_{fld} \tag{11}$$

where W_{pca} and W_{fld} represent principal component analysis and Fisher linear discrimination analysis [7]. W_{pca} is constructed by selecting principal components of S_t. We use a simple selection measure of principal components for W_{pca}. If the total number of components is less than $2 * c$, where c is the number of classes, then we keep all the components, otherwise we discard the smallest c ones like the original Fisherface method. This is an experimental measure. We find that after selecting frequency bands, the dimension of obtained feature vector is small, and the number of generated principal components is often lower than $2 * c$. In such a situation, if we discarded the smallest c components, then the number of remaining ones is lower than c and the recognition effect is often not ideal. Therefore, this measure is suitable for our proposed approach, which involves the method of selecting components in the original Fisherface method.

b)　Select the achieved discrimination vectors in the following way. Suppose that $W_{opt} = (\varphi_1, \varphi_2, \cdots, \varphi_M)$, where M is the number of vectors. The Fisher discrimination value of $\varphi_i (1 \le i \le M)$ is defined as follows:

$$F(\varphi_i) = \frac{\varphi_i^T \left(W_{pca}^T S_b W_{pca} \right) \varphi_i}{\varphi_i^T \left(W_{pca}^T S_w W_{pca} \right) \varphi_i}. \tag{12}$$

Select φ_i if $F(\varphi_i) > T_2$ and obtain the final discrimination transform matrix W. Similarly to T_1, the theoretical value of T_2 should be 1.0. However, its experimental value might not be completely consistent with its theoretical value. In the experiments, T_2 is set no more than 1. The reason is that extracting discrimination vectors in this step is after the selection of frequency bands in Step 1. In other words, we have carried out one selection procedure for using effective bands by setting T_1. Thus, for the generated discrimination vectors whose separability values are less than 1, they might have useful discrimination information for the classification. We need to make use of as many vectors as possible. Our experimental results will show that T_2 is set no more than 1 for all data.

Step 3)　For each sample z_1 in Z_1 and z_2 in Z_2, extract linear discrimination feature l_1 and l_2

$$l_1 = z_1 W \text{ and } l_2 = z_2 W. \tag{13}$$

Then, use the nearest neighbor classifier for classification. Here, the distance d between two training sample l_1 and test sample l_2, is defined by

$$d(l_1, l_2) = \|l_1 - l_2\|_2 \tag{14}$$

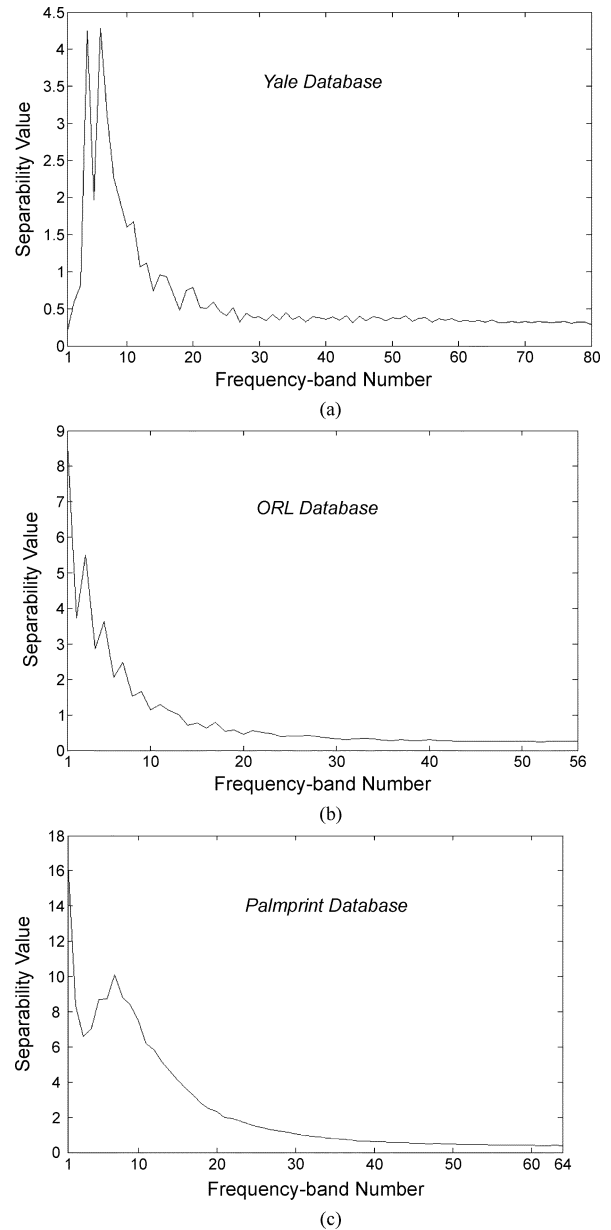where $\| \ \|_2$ denotes the Euclidean distance.

Fig. 4.　Demo of separability values of all frequency bands for different image data. (a) Yale face database. (b) ORL face database. (c) Palmprint database.

III. THEORETICAL ADVANTAGE ANALYSIS

In this section, we analyze the theoretical advantages of our approach.

A. Favorable Properties of DCT

The KL transform is an optimal transform for removing statistical correlation [20]. Of the discrete transforms, DCT best approaches the KL transform [21]. In other words, DCT has strong ability of removing correlation and compressing images. This is also illustrated by Hafed and Levine in face recognition [5]. They directly extracted DCT feature from facial images and achieved similar classification effect as the Eigenface method which is based on KL transform. Besides, DCT can be realized

TABLE I
SEPARABILITY VALUES OF DIFFERENT SUBIMAGES BY WAVELET DECOMPOSITION AND THE CORRESPONDING RECOGNITION RATES BY USING THE DISCRIMINANT WAVELETFACE METHOD ON ORL DATABASE

Classification performance	Different levels	Image size	Sub-images			
			Low-frequency	Horizontal edge	Vertical edge	Diagonal edge
Separability values	1	46*56	2.2430	0.4429	0.3649	0.2560
	2	23*28	2.6391	0.6837	0.5591	0.3145
	3	12*14	3.2727	1.2091	1.0479	0.4860
	4	6*7	4.2262	2.3663	1.5121	1.0043
Recognition rates (%)	1	46*56	N/A	N/A	N/A	N/A
	2	23*28	N/A	N/A	N/A	N/A
	3	12*14	94.5[14]	N/A	N/A	N/A
	4	6*7	95	84.5	71	58

by fast Fourier transform (FFT), while there is no fast realization algorithm for KL transform. Therefore, our approach sufficiently utilizes these favorable properties of DCT.

B. Precise Frequency-Band Selection

Another advantage of our approach is that it can precisely select appropriate frequency bands with favorable linear separability. Fig. 4(a)–(c) provides a demo of separability values of all bands for various image data Yale face database, ORL face database, and palmprint database, where the numbers of training samples per class for all data are five. From Fig. 4, an experiential rule can be obtained. The lower frequency bands generally have larger separability values, and there is no completely direct proportional relationship between the separability of a band and the band's level.

The discriminant waveletface method [14] extracts the third level low-frequency subimage of original image by using wavelet transform. According to the obtained experimental rule, three disadvantages exist for this method.

1) It cannot theoretically determine which level of subimage is most appropriate for extracting linear discrimination features.
2) Not all information in the low-frequency subimage is useful. Fig. 4(a) provides an effective illustration. The separability values of the first two frequency bands are smaller (less than 1). Corresponding to these bands, the related information of the subimage should be removed since it is useless to pattern classification.
3) The useful discriminant information of other subimages may be discarded. Table I shows the separability values of different subimages of wavelet decomposition calculated by the image separability judgment, where the types of subimages include low-frequency, horizontal edge, vertical edge and diagonal edge, and the levels of subimages are from one to four. This table also displays the recognition rates of different subimages by using the discriminant waveletface method, where the nearest neighbor classifier is adopted. For the edge subimages of the third and fourth levels, most of their separability values are more than 1.0. Moreover, the recognition effects of the fourth-level

edge subimages demonstrate that some useful discriminant information in them should not be discarded. Besides, from the fourth-level low-frequency subimages, we can obtain better recognition rate (95%) than that from the third-level low-frequency subimages (94.5%). This also illustrates the first disadvantage of the discriminant waveletface method.

Our approach can select all the appropriate frequency bands with favorable linear separability and discard the unuseful bands which do not satisfy the separability requirement. Furthermore, the DCT is a linear orthogonal transform and there is little correlation between different bands. Therefore, we can combine the selected bands to construct a feature set for the pattern classification. The experiments will demonstrate that this selection measure is helpful to improve the recognition effect and to reduce the dimension of feature space.

C. Operation Facility in Frequency Domain

The third advantage is that the operation facility of our approach. It can select the bands directly in the frequency-domain since the transformed resutls of DCT can be expressed by real number. However, if our approach is based on FT, we cannot directly select the bands in the frequency-domain. This is because the transformed resutls of FT are expressed by complex numbers. If we wish to evaluate the linear separabiltiy of frequency bands of FT, we must conduct inverse FT for the interesting bands. In other words, we must evaluate the separability of interesting bands in the space-domain of image. Obviously, this will increase the computational cost. Hence, our DCT-based approach can save the computing time than the potential FT-based method.

In the experiments, we have demonstrated that after selecting the frequency bands using our approach, the same total separabilty value can be achieved from the DCT frequency-domain images and the space-domain images generated using inverse DCT.

IV. EXPERIMENTAL RESULTS

This section will compare the experimental results of our approach with four conventional linear discrimination methods:
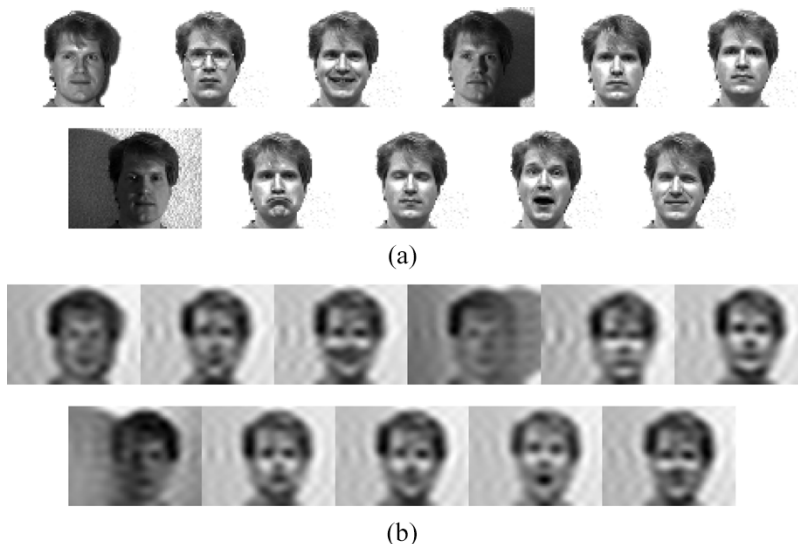
Fig. 5. Demo images from the Yale database. (a) Original facial images. (b) Processed images after band selection.

TABLE II
IMPLEMENT PROCEDURE OF OUR APPROACH FOR DIFFERENT IMAGE DATA

Implement procedure of our approach		Image data		
		Yale face database	ORL face database	Palmprint database
Parameter setting	T_1	0.8	2.0	2.0
	T_2	1.0	0.6	0.5
Important experimental results	Numbers of the selected frequency bands	3-13, 15-16	1-7	1-20
	Total separability of training set — Before band selection	1.540	1.994	5.014
	Total separability of training set — After band selection	1.955	3.914	8.458
	Feature vector dimension of selected bands	225	49	400
	Total number of principal components	74	49	210
	Number of classes	15	40	190
	Number of used principal components	59	49	210
	Extracted feature dimension	14	29	181

Eigenface, Fisherface, DLDA, and discriminant waveletface. All methods adopt the same classifier as our approach. The experiments are implemented on a Pentium 1.4-GHz computer (256 MB RAM) and programmed in the MATLAB language (Version 6.5).

A. Experiments With the Yale Face Database

The Yale face database (http://cvc.yale.edu) contains images with major variations, including changes in illumination, subjects wearing eyeglasses and different facial expressions. This database involves 165 frontal facial images, with 11 images of 15 individuals. The size of each image is 243×320 with 256 gray levels. To decrease computing time and simultaneously guarantee sufficient resolution, each image is scaled to an image size of 60×80. We use the full facial image without manually cutting out the background, which is different from the Fisherface method. Fig. 5(a) shows 11 sample images for one person. Fig. 5(b) shows the corresponding images processed by frequency band selection. We take the first five images of each person as the training samples and the remainder as the test samples. So, the numbers of training and test samples are 75 and 90. The related parameters in our approach for Yale database can be seen in Table II. The experimental values of T_1 and T_2 are set as 0.8 and 1.0, respectively. Not all the low frequency bands are selected. The first two and the fourteenth bands are discarded as they are useless to pattern discrimination. After band selection, the total separability value of the training sample set is improved by $0.415\ (= 1.955 - 1.540)$. The total number of principal components is 74 which is more than $2*c\ (= 2*15 = 30)$. According to the improved Fisherface method, the smallest 15 components are discarded. Hence, the first 59 components are used for achieving the discrimination transform. And 14 discrimination vectors are obtained. Note that the total number of components is equal to the rank of S_t defined in (8). S_t is used to solve W_{pca}.

TABLE III
COMPARISON OF CLASSIFICATION PERFORMANCE USING THE YALE DATABASE

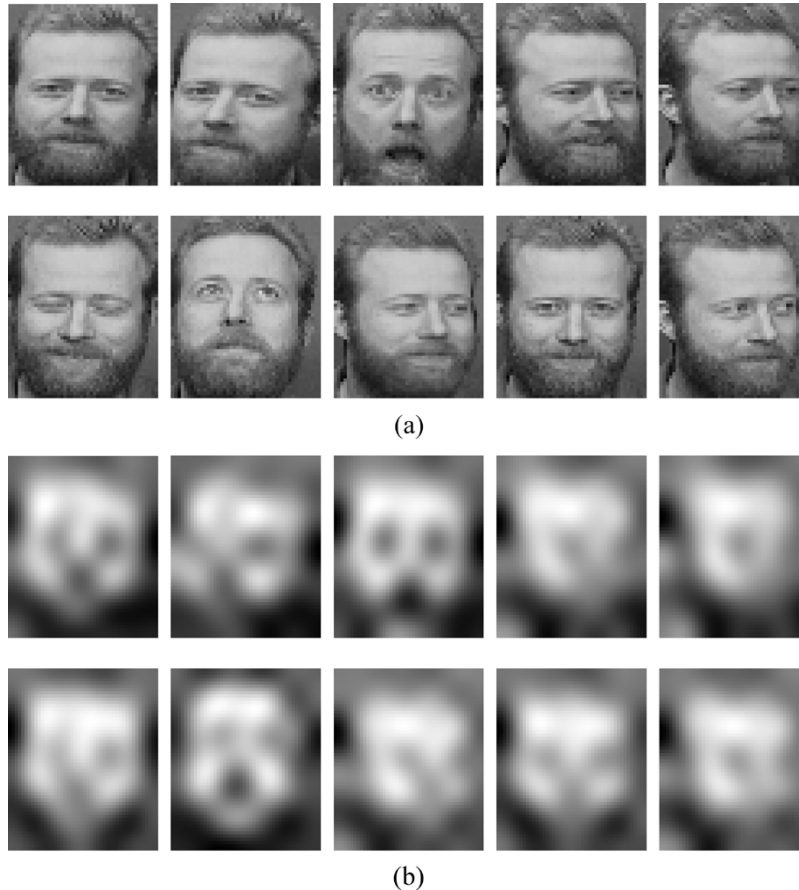| Methods	Our approach	Eigenface[6]	Fisherface[7]	DLDA[13]	Discriminant waveletface[14]
Recognition rates (%)	97.78	91.11	80	87.78	85.56
Extracted feature dimension	14	74	14	14	14
Training time (second)	14.8	14.3	14.9	13.1	15.2

(a)

(b)

Fig. 6. Demo images from the ORL database. (a) Original facial images. (b) Processed images after band selection.

A comparison of the classification performance of all the methods is provided in Table III. Our approach obtains the highest recognition rate. The maximum improvements in the recognition rate of our approach over those of Eigenface, Fisherface, DLDA and discriminant waveletface are 6.67%, 17.78%, 10%, and 12.22%, respectively. Notice that for the discriminant waveletface method, we take the fourth-level low-frequency subimages of the initial 60×80 images, that is, the subimage size is 4×5. With respect to the subimages of the first to the third levels, we cannot obtain the solution of discrimination transform since the within-class scatter matrix S_w in this method is singular. Our approach extracts the discriminative features with the same low dimension as other methods except for the Eigenface method. There is little difference in the training time of all methods.

Besides, there are three methods also using Yale face database to perform the experiments. In [25], [26], Jing *et al.* present some improvements on the linear discrimination analysis (LDA) and a generalized uncorrelated optimal discrim-

ination vectors (UODV) discrimination method, respectively. These two methods take first five images of each person as the training samples like our approach. The acquired recognition rates are 89.01% and 92.22%, which are less than the recognition result acquired by our approach, i.e. 97.78%. In [30], Dai and Yuen present a regularized discriminant analysis method for face recognition. Using Yale database, they obtained a mean recognition rate 97.5% on selecting arbitrary five images of each person as the training samples four times. It cannot compare with the recognition rate acquired by our approach because we use the first five images as the training samples.

B. Experiments With the ORL Face Database

The ORL database (http://www.cam-orl.co.uk) contains images varied in facial expressions, facial details, facial poses, and in scale. The database contains 400 facial images: ten images of 40 individuals. The size of each image is 92×112 with 256 gray levels. Each image is scaled to 46×56. Fig. 6(a) shows

TABLE IV
COMPARISON OF CLASSIFICATION PERFORMANCE USING THE ORL DATABASE

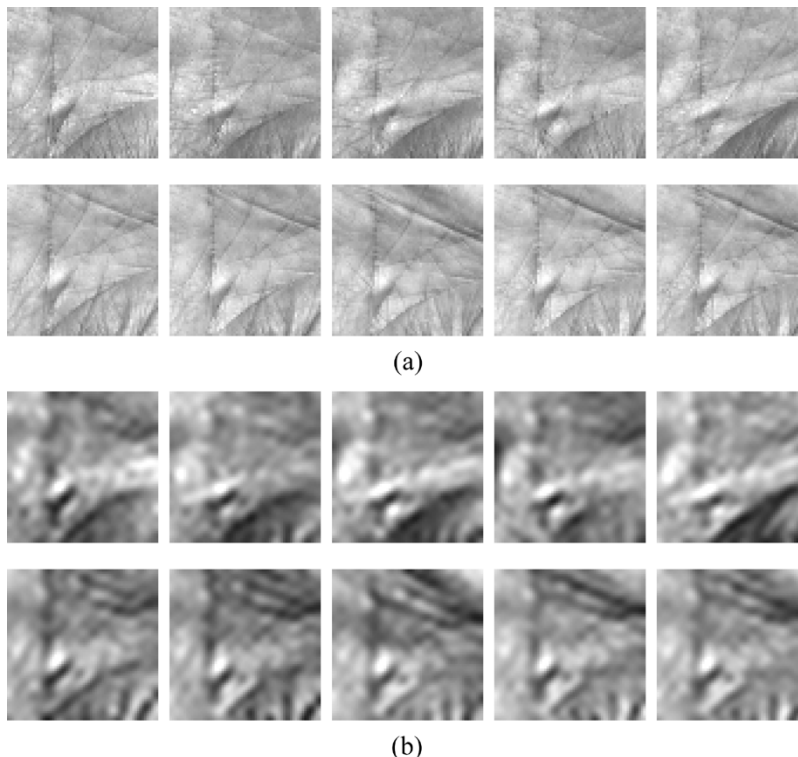| Methods	Our approach	Eigenface[6]	Fisherface[7]	DLDA[13]	Discriminant waveletface[14]
Recognition rates (%)	97.5	90	82.5	89	94.5[14]
Extracted feature dimension	29	199	39	39	39
Training time (second)	24.9	23.7	26.4	22.1	28.5

(a)

(b)

Fig. 7. Demo images from the palmprint database. (a) Original palmprint images. (b) Processed images after band selection.

ten sample images for one person. Fig. 6(b) shows the corresponding processed images by frequency band selection. We use the first five images of each person as the training samples and the remainder as the test samples. In other words, there is an equal number (200) of training and test samples. The related parameters in our approach for ORL database can also be seen in Table II. The experimental values of T_1 and T_2 are set as 2.0 and 0.6, respectively. Only a small part of the lowest frequency bands are selected, which are the first seven bands. The total separability value of the training sample set is remarkably improved by $1.92 (= 3.914 - 1.994)$ after band selection. The total number of principal components is 49 which is less than $2 * c (= 2 * 40 = 80)$. So, we do not discard any components. In the end, 29 discrimination vectors are obtained.

A comparison of the classification performance of all the methods is provided in Table IV. Our approach obtains the highest recognition rate and the lowest feature dimension. The maximum improvements in the recognition rate of our approach over those of Eigenface, Fisherface, DLDA and discriminant waveletface are 7.5%, 15%, 8.5%, and 3%, respectively. Compared with Fisherface, DLDA, and discriminant waveletface (which uses the second least number of features), our approach reduces the feature dimension by 25.64%. There is also little difference in the training time of all methods.

Some other methods also use ORL database. Ko *et al.* present a N-division output coding method for face recognition [28] and Yang *et al.* put forward an image PCA method [29]. These two methods take first five images of each person as the training samples like our approach. The acquired recognition rates are 93.5% and 95.5%, which are less than the recognition result acquired by our approach, i.e. 97.5%. In [30], Dai *et al.* present a regularized discriminant analysis method for face recognition. Using ORL database, they obtained a mean recognition rate 95.25% on selecting arbitrary five images of each person as the training samples four times. It cannot compare with the recognition rate acquired by our approach because we use the first five images as the training samples.

C. Experiments With the Palmprint Database

For reasons such as its accommodation of low-resolution imaging, its ability to operate on low-cost capture devices, and the ease with which the palm can be segmented, palmprint recognition has become an important complement to personal identification [22]. Wu *et al.* use the Fisherpalm method in palmprint recognition [27], which is very similar to the Fisherface method [7]. We collected palmprint images from 190 individuals using our self-designed capture device. The subjects

TABLE V
COMPARISON OF CLASSIFICATION PERFORMANCE USING THE PALMPRINT DATABASE

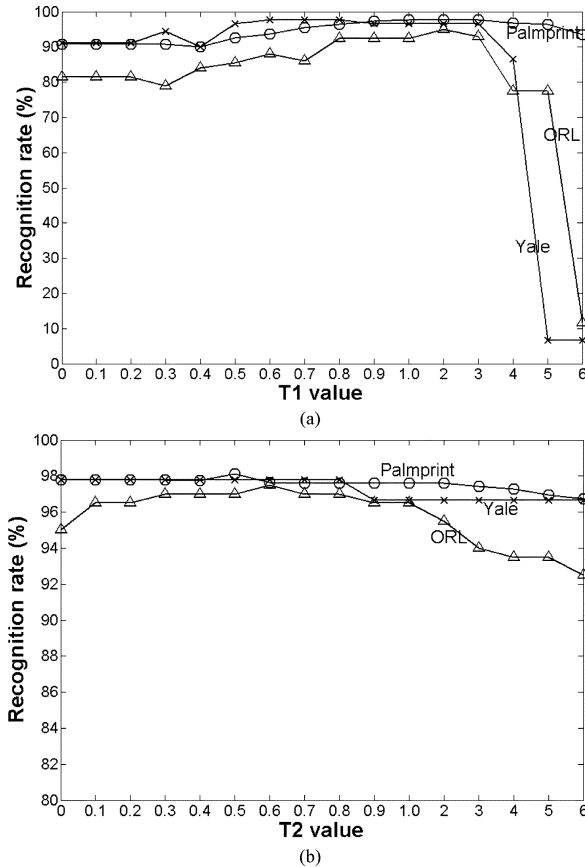| Methods	Our approach	Eigenface[6]	Fisherface[7]	DLDA[13]	Discriminant waveletface[14]
Recognition rates (%)	98.13	71.34	90.91	71	94.97
Extracted feature dimension	181	949	189	189	64
Training time (second)	196.6	204.3	323.9	161.9	172.7

Fig. 8. Recognition rates of our approach with different image data while (a) the value of T_1 is varied and (b) the value of T_2 is varied.

mainly consisted of student and staff volunteers from the Hong Kong Polytechnic University. Of the subjects in this database, 130 persons are male, approximately 87% of the subjects are younger than 30 years old, about 10% are aged between 30–50, and about 3% are older than 50. The palmprint images were collected on two separate occasions, at an interval of around two months. After finishing the first collection, we slightly changed the light source and adjusted the focus of the CCD camera so that the images collected on the first and second occasions might be regarded as being captured by two different palmprint devices. On each occasion, the subjects were asked to each provide eight palmprint images for the right hand. Thus, each person provides 16 images and our database contains a total of 3 040 images from 190 different palms. The size of all the original palmprint images is 384×284 pixels with 75 dpi resolution. Using the preprocessing approach in [23], the subimages with a fixed size (128×128) are extracted from the original images. In order to reduce the computational cost, each subimage is scaled to 64×64. We use these subimages to represent the original palmprint images and to conduct our experiments. Fig. 7(a) shows ten image samples of one person captured at different time. The first five were collected first collections and second five on the next occasion, the major changes being in illumination and position, including shift and rotation. Similar to the kinds of changes encountered in facial expressions, the image may also be slightly affected by the way the hand is posed, shrunk, or stretched. Fig. 7(b) shows the corresponding processed images by frequency band selection. We also use the first five images of each person as the training samples and the remainder as the test samples. So, the numbers of training and test samples are 950 and 2090. The related parameters in our approach for the palmprint database can be seen in Table II. The experimental values of T_1 and T_2 are set as 2.0 and 0.5, respectively. The first twenty low frequency bands are selected. After band selection, the total separability value of the training sample set is remarkably increased by $3.444 \, (= 8.458 - 5.014)$. The total number of principal components is 210 which is also less than $2 * c \, (= 2 * 190 = 380)$. So, we do not discard any components. And 181 discrimination vectors are obtained.

TABLE VI
ANALYSIS OF THE MEAN VALUES AND THE VARIANCES OF THE RECOGNITION RATES IN OUR APPROACH WHEN THE VALUE RANGES OF T_1 AND T_2 ARE [0.8, 3.0] AND [0.0, 2.0], RESPECTIVELY

Different thresholds	T_1			T_2		
	Different databases			Different databases		
	Yale	ORL	Palmprint	Yale	ORL	Palmprint
Mean recognition rates (%)	96.89	93.10	97.43	97.50	96.58	97.73
Variance	0.50	1.08	0.55	0.50	0.70	0.15

A comparison of the classification performance of all the methods is provided in Table V. Our approach obtains the highest recognition rate. The maximum improvements in the recognition rate of our approach over those of Eigenface, Fisherface, DLDA, and discriminant waveletface are 26.79%, 7.22%, 27.13%, and 3.16%, respectively. The second least number of features is acquired in our approach. We think that our approach makes a trade-off between obtaining the high recognition rate and reducing the dimension of feature space. It takes the third least training time in all methods, and there is no significant difference in the time of the fastest method (DLDA) and our approach.

D. Analysis of Threshold Setting

We perform some analysis for setting the experimental values of T_1 and T_2. Fig. 8 illustrates the recognition rates of three image databases while the values of T_1 and T_2 are varied. From Fig. 8(a) we find that with respect to T_1, the appropriate value ranges for Yale, ORL and palmprint databases are [0.5, 3.0], [0.8, 3.0], and [0.7, 5.0], respectively. That is in each range, all the recognition rates are near to the maximal rate. Hence, an appropriate range for both data is [0.8, 3.0]. From Fig. 8(b) we find that with respect to T_2, the appropriate value ranges for Yale, ORL and palmprint databases are [0.0, 6.0], [0.0, 2.0] and [0.0, 3.0], respectively. Hence, an appropriate range for both data is [0.0, 2.0]. Table VI shows an analysis of the mean values and the variances of the recognition rates where the value ranges of T_1 and T_2 are [0.8, 3.0] and [0.0, 2.0]. The variances are much smaller than the mean values, especially for T_2. That is, in these ranges the recognition effect of our approach is robust.

V. CONCLUSION

This paper presents a novel face and palmprint recognition approach based on DCT and linear discrimination technique. A 2-D separability judgment is used to select appropriate DCT frequency bands with favorable linear separability. And an improved Fisherface method is then applied to extract linear discriminative features from the selected bands. The detailed analysis shows the theoretical advantages of our approach over other frequency-domain transform techniques and state-of-the-art linear discrimination methods. The practicality of our approach as an image recognition approach is well evidenced in the experimental results, where different image data including two face databases and a palmprint database are used. Our approach can significantly improve image recognition effect. In contrast with four conventional discrimination methods (Eigenface, Fisherface, DLDA, and discriminant waveletface), it improves the average recognition rates of all data by 13.65%, 13.33%, 15.21%, and 6.13%, respectively. Besides, our approach can reduce the dimension of feature space and cost little computing time.

REFERENCES

[1] W. Li, D. Zhang, and Z. Xu, "Palmprint identification by Fourier transform," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 16, no. 4, pp. 417–432, 2002.
[2] J. H. Lai, P. C. Yuen, and G. C. Feng, "Face recognition using holistic Fourier invariant features," *Pattern Recognit.*, vol. 34, no. 1, pp. 95–109, 2001.
[3] R. Chellappa, C. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proc. IEEE*, vol. 83, pp. 705–740, May 1995.
[4] Y. Tian, T. N. Tan, Y. H. Wang, and Y. C. Fang, "Do singular values contain adequate information for face recognition?," *Pattern Recognit.*, vol. 36, no. 3, pp. 649–655, 2003.
[5] Z. M. Hafed and M. D. Levine, "Face recognition using the discrete cosine transform," *Int. J. Comput. Vis.*, vol. 43, no. 3, pp. 167–188, 2001.
[6] M. Turk and A. Pentland, "Eigenfaces for recognition," *Int. J. Cog. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
[7] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherface: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 19, pp. 711–720, July 1997.
[8] A. M. Martinez and A. C. Kak, "PCA versus LDA," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 23, pp. 228–233, Feb. 2001.
[9] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Ann. Eugenics*, vol. 7, pp. 178–188, 1936.
[10] C. Liu and H. Wechsler, "Robust coding scheme for indexing and retrieval from large face databases," *IEEE Trans. Image Processing*, vol. 9, pp. 132–137, Jan. 2000.
[11] ——, "A shape- and texture-based enhanced Fisher classifier for face recognition," *IEEE Trans. Image Processing*, vol. 10, pp. 598–608, Apr. 2001.
[12] L. Chen, H. M. Liao, M. Ko, J. Lin, and G. Yu, "A new LDA-based face recognition system which can solve the small sample size problem," *Pattern Recognit.*, vol. 33, no. 10, pp. 1713–1726, 2000.
[13] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data with application to face recognition," *Pattern Recognit.*, vol. 34, no. 12, pp. 2067–2070, 2001.
[14] J. T. Chien and C. C. Wu, "Discriminant waveletfaces and nearest feature classifiers for face recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, pp. 1644–1649, Dec. 2002.
[15] D. Zhang, H. Peng, J. Zhou, and S. K. Pal, "A novel face recognition system using hybrid neural and dual eighefaces methods," *IEEE Trans. Syst., Man, Cybern. A*, vol. 32, pp. 787–793, Nov. 2002.
[16] W. Malina, "Two-parameter Fisher criterion," *IEEE Trans. Syst., Man, Cybern. B*, vol. 31, pp. 629–636, Aug. 2001.

[17] T. Cooke, "Two variations on Fisher's linear discrimination for pattern recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 24, pp. 268–273, Feb. 2002.

[18] M. H. Yang, "Kernel eigenfaces vs. kernel fisherfaces: Face recognition using kernel methods," in *IEEE Proc. 5th Int. Conf. Automatic Face Gesture Recognition (RGR)*, pp. 215–220.

[19] M. S. Bartlett, J. R. Movellan, and T. J. Sejnowski, "Face recognition by independent component analysis," *IEEE Trans. Neural Network*, vol. 13, pp. 1450–1464, Nov. 2002.

[20] A. Rosenfeld and A. Kak, *Digital Picture Processing*. New York: Academic, 1976.

[21] A. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ: Prentice–Hall, 1989.

[22] D. Zhang, *Automated Biometrics—Technologies and Systems*. Norwell, MA: Kluwer, 2000.

[23] D. Zhang, W. K. Kong, J. You, and M. Wong, "On-line palmprint identification," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 25, pp. 1041–1050, Sept. 2003.

[24] X. Y. Jing, D. Zhang, and J. Y. Yang, "Face recognition based on a group decision-making combination approach," *Pattern Recognit.*, vol. 36, no. 7, pp. 1675–1678, 2003.

[25] X. Y. Jing, D. Zhang, and Y. F. Yao, "Improvements on the linear discrimination technique with application to face recognition," *Pattern Recognit. Lett.*, vol. 24, no. 15, pp. 2695–2701, 2003.

[26] X. Y. Jing, D. Zhang, and Z. Jin, "UODV: Improved algorithm and generalized theory," *Pattern Recognit.*, vol. 36, no. 11, pp. 2593–2602, 2003.

[27] X. Wu, D. Zhang, and K. Wang, "Fisherpalms based on palmprint recognition," *Pattern Recognit. Lett.*, vol. 24, no. 15, pp. 2829–2938, 2003.

[28] J. Ko and H. Byun, "N-division output coding method applied to face recognition," *Pattern Recognit. Lett.*, vol. 24, no. 16, pp. 3115–3123, 2003.

[29] J. Yang and J. Y. Yang, "From image vector to matrix: A straightforward image projection technique—IMPCA vs. PCA," *Pattern Recognit.*, vol. 35, no. 9, pp. 1997–1999, 2002.

[30] D. Q. Dai and P. C. Yuen, "Regularized discriminant analysis and its application to face recognition," *Pattern Recognit.*, vol. 36, no. 3, pp. 845–847, 2003.

Xiao-Yuan Jing received the M.S. and Ph.D. degrees in pattern recognition from Nanjing University of Science and Technology, Nanjing, China.

He was an Associate Professor at the Chinese Academy of Sciences, Beijing, China, and a Visiting Scholar at Hong Kong Polytechnic University, Hong Kong, China, and Hong Kong Baptist University, Hong Kong. His current interests include pattern recognition, image processing, neural networks, machine learning, and artificial intelligence.
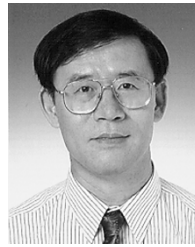
David Zhang received the B.S. degree in computer science from Peking University, Peking, China, in 1974, the M.Sc. and Ph.D. degrees in computer science and engineering from the Harbin Institute of Technology (HIT), Harbin, China, in 1983 and 1985, and the Ph.D. degree in electrical and computer engineering from the University of Waterloo, ON, Canada, in 1994.

From 1986 to 1988 he was a PostDoctoral Fellow at Tsinghua University, Beijing, China, and then an Associate Professor at the Academia Sinica, Beijing. He is currently at the Hong Kong Polytechnic University, Kowloon, China, where he is the Founding Director of the Biometrics Technology Center (UGC/CRC) a body supported by the Hong Kong SAR Government. He also serves as Adjunct Professor in Tsinghua University, Shanghai Jiao Tong University, Harbin Institute of Technology, and the University of Waterloo. His research interests include automated biometrics-based authentication, pattern recognition, and biometric technology and systems. He is the Founder and Editor-in-Chief of *International Journal of Image and Graphics* (IJIG), Book Editor for the Kluwer International Series on Biometrics (KISB), Associate Editor of more than ten international journals including *Pattern Recognition*, and author of more than 130 journal papers, 20 book chapters, and ten books. He holds a number of patents in both the U.S.A. and China and is a current Croucher Senior Research Fellow.

Dr. Zhang, in 1984, won the National Scientific Council of China's third prize for his Fingerprint Recognition System, and in 1986, his Real-Time Remote Sensing System took the Council's first prize. In 2002, his Palmprint Identification System won a Silver Medal at the Seoul International Invention Fair, following that in 2003 by taking a Special Gold Award, a Gold Medal, and a Hong Kong Industry Award. He is Associate Editor for IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS–PART A and TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS–PART C. He was Program Chair for the First International Conference on Biometrics Authentication (ICBA).