Pattern Recognition Letters 31 (2010) 2337–2344

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Directional binary code with application to PolyU near-infrared face database

Baochang Zhang [b], Lei Zhang [a,*], David Zhang [a], Linlin Shen [c]

[a] Biometrics Research Center, Dept. of Computing, The Hong Kong Polytechnic University, Hong Kong, Hong Kong
[b] National Key Laboratory of Science and Technology on Integrated Control Technology, School of Automation Science and Electrical Engineering, Beihang University, Beijing, China
[c] School of Computer Science & Software Engineering, Shenzhen University, China

ARTICLE INFO

Article history:
Received 21 August 2009
Available online 23 July 2010
Communicated by G. Borgefors

Keywords:
Near-infrared face recognition
Face database
Feature extraction

ABSTRACT

This paper introduces the establishment of PolyU near-infrared face database (PolyU-NIRFD) and presents a new coding scheme for face recognition. The PolyU-NIRFD contains images from 350 subjects, each contributing about 100 samples with variations of pose, expression, focus, scale, time, etc. In total, 35,000 samples were collected in the database. The PolyU-NIRFD provides a platform for researchers to develop and evaluate various near-infrared face recognition techniques under large scale, controlled and uncontrolled conditions. A new coding scheme, namely directional binary code (DBC), is then proposed for near-infrared face recognition. Finally, we provide three protocols to evaluate and compare the proposed DBC method with baseline face recognition methods, including Gabor based Eigenface, Fisherface and LBP (local binary pattern) on the PolyU-NIRFD database. In addition, we also conduct experiments on the visible light band FERET database to further validate the proposed DBC scheme.

1. Introduction

Face recognition (FR) is a promising technology for automated personal authentication and it has a great potential in applications of public security, video surveillance, access control, forensics, etc. (Chellappa et al., 1995; Zhao et al., 2003). Meanwhile, FR is one of the most active topics in the filed of computer vision, and several large-scale face databases (Gao et al., 2008; Enrique et al., 2003; Sim et al., 2003; Messer et al., 1999; Martinez and Benavente, 1998; Georghiades et al., 2001; Lee et al., 2005) are publicly available for evaluating and comparing various FR methods. Generally speaking, FR in visible spectrum has been extensively studied because it is convenient to implement in various environments and has a wide range of applications (Chellappa et al., 1995; Zhao et al., 2003; Turk and Pentland, 1991; Belhumeur et al., 1997). Many FR algorithms have been proposed (Chellappa et al., 1995; Zhao et al., 2003), and the large-scale face databases play an important role in evaluating and developing FR algorithms.

Face recognition technology (FERET) (Phillips et al., 2000), face recognition vendor test (FRVT) (Phillips et al., 2003), and face recognition grand challenge (FRGC) (Philips et al., 2005) have pioneered both evaluation protocols and database construction. FRGC is more changeling than FERET and FRVT, as it contains more uncontrolled variations and 3D images in its database. For example, in the most challenging set of the FRGC v2 (Exp#4), the training set contains 10,776 images from 222 subjects, while the query and target sets contain 8014 and 16,028 images, respectively. The 3D training set of FRGC consists of controlled and uncontrolled still images from 943 subject sessions, while the validation partition of FRGC contains images from 466 subjects collected in 4007 subject sessions. Other publicly available face databases include the CAS-PEAL (Gao et al., 2008), BANCA (Enrique et al., 2003), CMU PIE (Sim et al., 2003), XM2VTSDB (Messer et al., 1999), AR (Martinez and Benavente, 1998), Yale (Georghiades et al., 2001; Lee et al., 2005), etc. These face databases in visible spectrum provide a good evaluation platform for various FR techniques, and in return they greatly facilitate the development of new FR methods.

In visual face recognition, the performance suffers from the lighting variations. Traditional methods are mostly based on the Lambertian model (Sim et al., 2003), which is however too simply to describe the real face surface under various illuminations. Some illumination invariant FR methods, such as the method based on thermal infrared images (Kong et al., 2007), have been developed. Recently, active near-infrared (NIR) FR was proposed to deal with the illumination variations in different environments, and NIR-based FR has shown promising performance in real applications. NIR-based FR uses imaging sensors in invisible spectral bands to reduce the affection of ambient light. The NIR band is between the visible light band and the thermal infrared band, and it has advantages over both visible light and thermal infrared. Firstly, since NIR light can be reflected by objects, it can serve as an active illumination source, in contrast to thermal infrared. Secondly, it is invisible, making active NIR illumination unobtrusive. Thirdly, unlike thermal infrared, NIR can easily penetrate glasses (Zou et al., 2007).

* Corresponding author. Tel.: +852 27767355; fax: +852 27740842.
E-mail address: cslzhang@comp.polyu.edu.hk (L. Zhang).

0167-8655/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2010.07.006

Several NIR FR systems have been proposed. Zou et al. proposed to use active NIR light to localize face areas in the images and then recognize faces (Zou et al., 2005). Li et al. extracted the local binary pattern feature and used Fisher analysis for NIR-based FR, and they developed a complete NIR FR system which can perform face detection, eye localization and face identification (Li et al., 2007). Pan et al. (2003) proposed an NIR-based FR system which captures face images in wavelength of 0.7–1.0 μm. Some other works using NIR images for FR can be found in (Dowdall et al., 2003; Georghiades et al., 2001). Zou et al. (2005) have shown that FR in NIR band has better performance than that in the visible band, and this was also validated in Li's work (Li et al., 2007). However, so far there is not a large-scale NIR face database which is publicly available. There is a high demand to construct an open NIR FR database, on which the researchers can test and compare their algorithms. In this paper, we will introduce such a database we constructed in the Hong Kong Polytechnic University, namely the PolyU near-infrared face database (PolyU-NIRFD).

The face images in PolyU-NIRFD were collected from 350 subjects, each subject providing about 100 samples. The sample images involve various variations of expression, pose, scale, focus, time, etc. To evaluate the performance of different FR methods on the PolyU-NIRFD, we provide three test protocols, including the partition strategy of the training, gallery and probe sets and the baseline evaluation schemes. A new feature extraction method, namely directional binary code (DBC), is also proposed for NIR face image recognition. In DBC, the directional edge features are extracted and used to represent NIR face images, which are generally smoother than face images in visible band. The baseline algorithms we used for comparison are Eigenface, Fisherface, local binary pattern (LBP) (Ojala et al., 2002) and their Gabor filtering enhanced versions, which are well-known and representative methods in the field of FR. Considering that the Gabor filtering can improve significantly the FR accuracy, we also couple Gabor filtering with DBC to examine its performance on PolyU-NIRFD.

The rest of the paper is organized as follows. Section 2 introduces the developed NIR face acquisition system and the establishment of PolyU-NIRFD. In Section 3, the DBC feature extraction method is proposed for NIR FR. Section 4 presents extensive experiments using three protocols to evaluate the performance of various FR methods, including Gabor-Eigenface, Gabor-Fisherface, Gabor-LBP and DBC, on the PolyU-NIRFD. We also provide another experiment on the visible light band FERET database to further validate the performance of DBC. Conclusions are drawn in Section 5.

2. PolyU-NIRFD construction

Different from the FR in visible light band, which can simply use a common camera to capture face images, FR in NIR band needs some additional hardware and special system design for image acquisition. This section describes the NIR image acquisition system and the collection of the PolyU-NIRFD.

2.1. Near-infrared face image acquisition

The hardware of our NIR face image acquisition system consists of a camera, an LED (light emitting diode) light source, a filter, a frame grabber card and a computer. A snapshot of the constructed imaging system is shown in Fig. 1. The camera used is a JAI camera, which is sensitive to NIR band. The active light source is in the NIR spectrum between 780 and 1100 nm and it is mounted on the camera. The peak wavelength is 850 nm, and it lies in the invisible and reflective light range of the electromagnetic spectrum. An NIR LED array is used as the active radiation sources, and it is strong enough for indoor use. The LEDs are arranged in a circle and they are mounted on the camera to make the illumination on the face is as homogeneous as possible. The strength of the total LED lighting is adjusted to ensure a good quality of the NIR face images when the camera-face distance is between 80 and 120 cm, which is convenient for the users. When mounted on the camera, the LEDs are approximately coaxial to the imaging direction and thus provide the best possible straight frontal lighting. Although NIR is invisible to the naked eyes, many CCD cameras have sufficient response to the NIR spectrum. The filter we used in the device is used to cut off the visible light, whose spectrum is shorter (780 nm, visible light). For the convenience of data collection, we put the imaging device into a black box of 19 cm width, 19 cm long, and 20 cm high, as shown in Fig. 1.

2.2. PolyU-NIRFD

By using the self-designed data acquisition device described in Section 2.1, we collected NIR face images from 350 subjects. During the recording, the subject was first asked to sit in front of the camera, and the normal frontal face images of him/her were collected. Then the subject was asked to make expression and pose changes and the corresponding images were collected. To collect face images with scale variations, we asked the subjects to move near to or away from the camera in a certain range. At last, to collect face images with time variations, samples from 15 subjects were

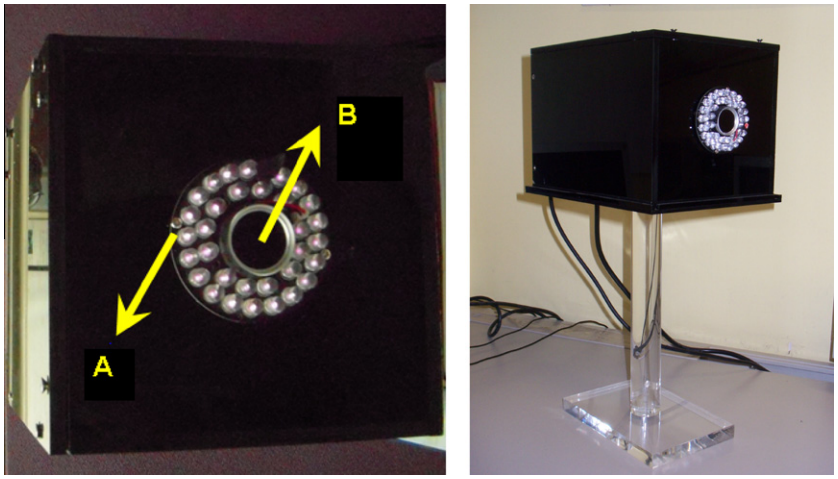

Fig. 1. The NIR face image acquisition device. 'A' is the NIR LED light source, and 'B' is the NIR sensitive camera with an NIR filter.

collected at two different times with an interval of more than 2 months. In each recording, we collected about 100 images from each subject, and in total 35,000 images were collected in the PolyU-NIRFD database. The sample images in the PolyU-NIRFD are labeled as 'NN_xxxxxx_S_D_****', where "NN" represents the prefix of the label string, 'S' represents the Gender information, 'xxxxxx' indicates the ID serial number of the subject, 'D' denotes the place where the image was captured, and '****' is the index of the face image. For example, "NN_200001_F_B_024" means that the 24th image is from a Female subject collected in Beihang University. Fig. 2 shows some face images of a subject with variations of expression, pose and scale. Fig. 3 shows the images of some subjects which were taken in different times.

To evaluate the performance of different methods on the PolyU-NIRFD, we design three types of experiments, each of which contains a training set, a target (Gallery) set and a query (Probe) set. In Exp#1, the used images include frontal face images as well as images with expression variations, scale changes (include blurring), time difference, etc. In this experiment, the sizes of the training set, target set and query set are 419, 574 and 2763, respectively. In Exp#2, we add more faces captured in uncontrolled conditions to make the test more challenging. The sizes of the training set, target set and query set are 1876, 1159, and 4747, respectively. In Exp#3, we focus on the images with high pose variations and exclude the images with expression, scale and time variations. The sizes of the training set, target set and query set are 576, 951, and 3648, respectively. In next section, we will present a new face feature coding algorithm, and then in Section 4 the three types of experiments will be conducted by different methods.

3. Directional binary code

3.1. Feature extraction by directional binary code (DBC)

In this section we present a simple yet efficient method for NIR face image feature extraction and representation, namely directional binary code (DBC). The DBC is proposed to encode the directional edge information in a neighborhood. Given an image I, we denote its first-order derivatives along 0°, 45°, 90° and 135° directions as $I'_{\alpha,d}$, where $\alpha = 0°, 45°, 90°$ and 135°, and d is the distance between the given point and its neighboring point. For example, in Fig. 4 the distance between the center point and its four directional neighbors is 1, i.e. $d = 1$ in four directions. Let $z_{i,j}$ be a point in I, then the four directional derivatives at $z_{i,j}$ are

$$\begin{cases} I'_{0°,d}(z_{i,j}) = I(z_{i,j}) - I(z_{i,j-d}) \\ I'_{45°,d}(z_{i,j}) = I(z_{i,j}) - I(z_{i-d,j+d}) \\ I'_{90°,d}(z_{i,j}) = I(z_{i,j}) - I(z_{i-d,j}) \\ I'_{135°,d}(z_{i,j}) = I(z_{i,j}) - I(z_{i-d,j-d}) \end{cases} \tag{1}$$

A thresholding function, $f(I'_{\alpha,d}(z))$, is applied to the four directional derivatives to output a binary code in the given direction:

$$f(I'_{\alpha,d}(z)) = \begin{cases} 1, & \text{if } I'_{\alpha,d}(z) \geqslant 0 \\ 0, & \text{if } I'_{\alpha,d}(z) < 0 \end{cases} \tag{2}$$

With Eq. (2), the DBC in direction α is defined as:

$$\begin{aligned} DBC_{\alpha,d}(z_{x,y}) = \Big\{ & f(I'_{\alpha,d}(z_{x,y}); f(I'_{\alpha,d}(z_{x,y-d}); f(I'_{\alpha,d}(z_{x-d,y-d}); f(I'_{\alpha,d}(z_{x-d,y}); \\ & f(I'_{\alpha,d}(z_{x-d,y+d}); f(I'_{\alpha,d}(z_{x,y+d}); f(I'_{\alpha,d}(z_{x+d,y+d}); f(I'_{\alpha,d}(z_{x+d,y}); \\ & f(I'_{\alpha,d}(z_{x+d,y-d}) \Big\} \end{aligned} \tag{3}$$

Fig. 5 shows an example to illustrate the calculation of DBC pattern along 0° at the central position.

The DBC can reflect the image local feature, and we model the global distribution of DBC by its histogram over the whole face image. We partition the face image into L regions, represented by $R_1, R_2, \ldots, R_L$. Denote by $H_{DBC_{\alpha,d}}(R_i)$ the histogram of $DBC_{\alpha,d}$ in region R_i. The shape of the regions we used in the experiments is rectangle, as shown in Fig. 6. The ordering is from left to right and top to bottom. The histogram of DBC in the whole image is defined as

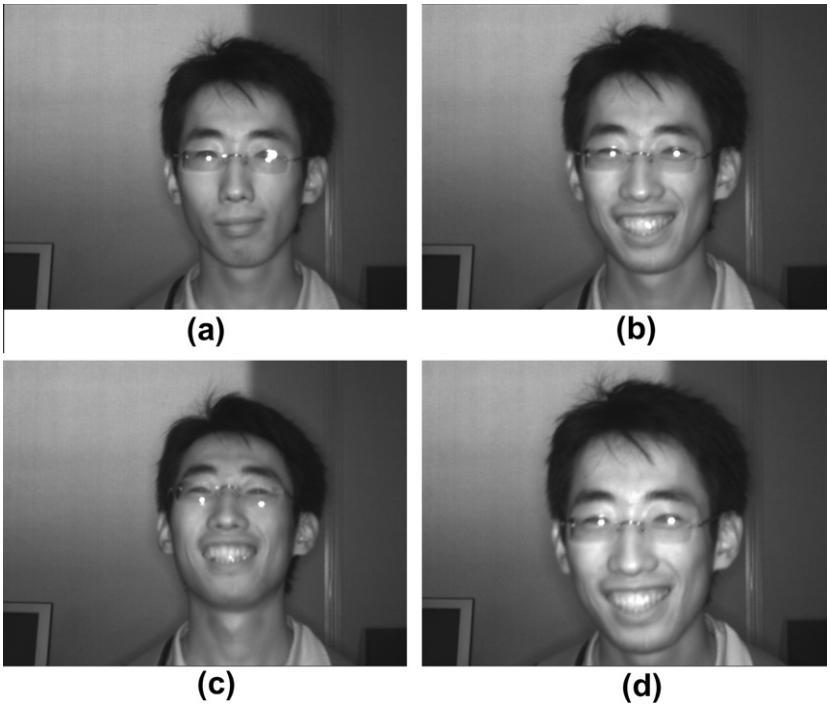

Fig. 2. Sample NIR face images of a subject. (a) Normal face image; and images with (b) expression variation; (c) pose variation and (d) scaling variation.

Fig. 3. Sample NIR face images captured in more than 2 months.

$I_{x-1,y-1}$	$I_{x,y-1}$	$I_{x+1,y-1}$
$I_{x-1,y}$	$I_{x,y}$	$I_{x+1,y}$
$I_{x-1,y+1}$	$I_{x,y+1}$	$I_{x+1,y+1}$

Fig. 4. A 3 × 3 neighborhood centered on $I_{x,y}$.

$$HDBC = \{H_{DBC_{\alpha,d}}(R_i)|i = 1,\ldots,L\} \quad (4)$$

Note that the regions do not have to be rectangular or of the same size. For example, spatial histograms can be extracted from circular regions with different radiuses.

The proposed DBC is different from the well-known LBP (Ahonen et al., 2006). The DBC captures the spatial relationship between any pair of neighborhood pixels in a local region along a given direction, while LBP considers the relationship between a given pixel and its surrounding neighbors. Therefore, DBC captures more spatial information than LBP. It has also been validated that the directional features are very valuable for object recognition (Liu and Wechsler, 2002; Gao and Leung, 2002). Fig. 7 shows an example to illustrate the difference between LBP and DBC. It can be seen that DBC can extract more edge information than LBP. The increment sign correlation (IST) method (Kanekoa et al., 2002) has some similar idea to the proposed DBC, but they are different as DBC focuses on local pattern features, while IST is based on the average evaluation of incremental tendency of brightness in adjacent pixels and normalized to be a binary distribution or a Gaussian distribution for a large image size through statistical analysis (Kanekoa et al., 2002).

3.2. Gabor based DBC (GDBC)

In image processing and pattern recognition, Gabor filtering is widely used to enhance the image features by using a set of Gabor filters, which can model the receptive field profiles of cortical simple cells (Wiskott et al., 1997; Liu and Wechsler, 2002). The Gabor filters can enhance salient visual features in an image, such as directional spatial-frequency characteristics, because they can selectively enhance features in certain scales and orientations. The Gabor filtering based Eigenface, Fisherface, and LBP schemes have shown significant improvement over the original versions (i.e. image based methods) (Liu and Wechsler, 2002; Zhang et al., 2005; Gao et al., 2008; Hong et al., 1998). In these methods, Gabor features are used as one kind of preprocessing tool, which are robust to certain degree of variations of illumination, pose, etc. However, it should be noted that Gabor filtering only enhances the features but it does not code the discriminative features. Therefore, after Gabor filtering, a feature extraction or feature coding process is required. In this section, we investigate the feasibility and effectiveness of Gabor based DBC features. Like in Gabor based Eigenface, Fisherface and LBP methods, where PCA, LDA and LBP are used as the discriminative feature extractors, in the proposed method the DBC is used to code the Gabor enhanced features.

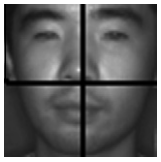

Fig. 6. The division of the face image.

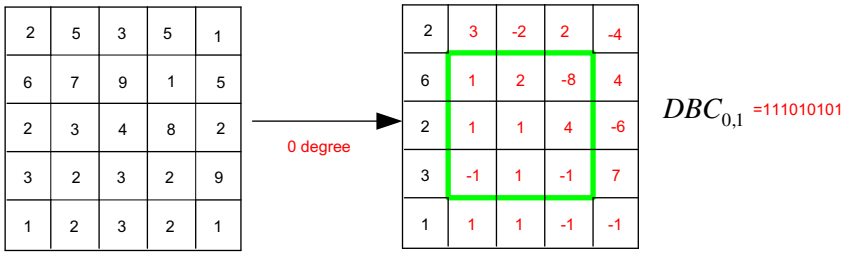

Fig. 5. An example of DBC along 0°.

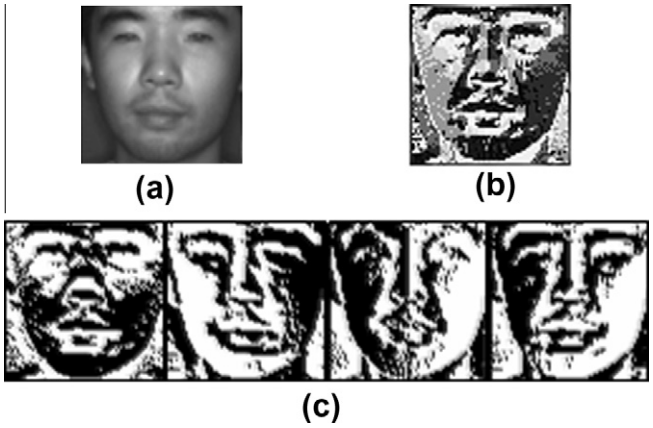

Fig. 7. Example of LBP and DBC feature maps: (a) is the original image; (b) is the LBP feature map; and (c) shows the DBC feature maps along 0°, 45°, 90° and 135° directions.

The Gabor filters are defined as follows:

$$\psi_{u,v}(\boldsymbol{z}) = \frac{\|\boldsymbol{k}_{u,v}\|^2}{\sigma^2} e^{(\|-\boldsymbol{k}_{u,v}\|^2\|\boldsymbol{z}\|^2/2\sigma^2)}[e^{i\boldsymbol{k}_{u,v}\boldsymbol{z}} - e^{-\sigma^2/2}] \tag{5}$$

where $\boldsymbol{z} = \begin{pmatrix} y \\ x \end{pmatrix}$, $\boldsymbol{k}_{u,v} = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_v \cos\phi_u \\ k_v \sin\phi_u \end{pmatrix}$, $k_v = \pi/2^{v/2}$, $\phi_u = u\frac{\pi}{8}$, $v = 0, \ldots, v_{\max} - 1$, $u = 0, \ldots, u_{\max} - 1$. v is the frequency, u is the orientation, $v_{\max} = 5$, $u_{\max} = 8$ and $\sigma = 2\pi$. For 2D images, the filtering is often performed along different directions to enhance the directional features. As a preprocessing procedure, Gabor filtering can effectively suppress noise and enhance the image salient features in given directions. In contrast, DBC is a kind of local operator to code the local pattern but it does not enhance the local features. Therefore, we can use DBC to code the Gabor enhanced local features for better face recognition results. In other words, Gabor filtering and DBC play different and complementary roles in the whole scheme.

Denote by $G_{u,v}(z_{i,j})$ the Gabor features at pixel $z_{i,j}$ of an image, where u and v are the orientation and scale of the kernel, respectively. Similar to Eq. (1), its directional derivatives along 0°, 45°, 90° and 135° are computed as

$$\begin{cases} G'_{u,v,0^\circ,d}(z_{i,j}) = G_{u,v}(z_{i,j}) - G_{u,v}(z_{i,j-d}) \\ G'_{u,v,45^\circ,d}(z_{i,j}) = G_{u,v}(z_{i,j}) - G_{u,v}(z_{i-d,j+d}) \\ G'_{u,v,90^\circ,d}(z_{i,j}) = G_{u,v}(z_{i,j}) - G_{u,v}(z_{i-d,j}) \\ G'_{u,v,135^\circ,d}(z_{i,j}) = G_{u,v}(z_{i,j}) - G_{u,v}(z_{i-d,j-d}) \end{cases} \tag{6}$$

We denote by $G'^{,\mathrm{Re}}_{u,v,\alpha,d}(z_{i,j})$ and $G'^{,\mathrm{Im}}_{u,v,\alpha,d}(z_{i,j})$, respectively, the real and imaginary parts of $G'_{u,v,\alpha,d}(z_{i,j})$. The Gabor based DBC (GDBC) in direction α and at position $z_{i,j}$ is then defined as

$$\begin{aligned} GDBC_{u,v,\alpha,d}(z_{i,j}) = \Big\{ & f(G'^{,\mathrm{Re}}_{u,v,\alpha,d}(z_{i,j+d})), f(G'^{,\mathrm{Im}}_{u,v,\alpha,d}(z_{i-d,j+d})), \\ & f(G'^{,\mathrm{Re}}_{u,v,\alpha,d}(z_{i+d,j})), f(G'^{,\mathrm{Im}}_{u,v,\alpha,d}(z_{i+d,j+d})), f(G'^{,\mathrm{Re}}_{u,v,\alpha,d}(z_{i,j-d})), \\ & f(G'^{,\mathrm{Im}}_{u,v,\alpha,d}(z_{i+d,j-d})), f(G'^{,\mathrm{Re}}_{u,v,\alpha,d}(z_{i-d,j})), f(G'^{,\mathrm{Im}}_{u,v,\alpha,d}(z_{i-d,j-d})) \Big\} \end{aligned} \tag{7}$$

where $f(\cdot,\cdot)$ is the thresholding function in Eq. (2). Please note that the alternating use of real and imaginary parts in the definition of GDBC is to reduce the length of the code while involving the real and imaginary information. It lies in the fact that the Gabor features extracted in neighborhood have high redundancy so that there is no necessary to keep the features at all positions. An alternative use of the real and imaginary parts can reduce the feature length in half.

After partition the face image into L rectangular regions as shown in Fig. 6, the histogram of GDBC at region R_i is defined as

$$HGDBC = \Big\{ H_{GDBC_{u,v,\alpha,d}}(R_i) | u = 0, \ldots, 7;\ v = 0, \ldots, 4;\ i = 1, \ldots, L; \\ \alpha = 0^\circ, 45^\circ, 90^\circ, 135^\circ \Big\} \tag{8}$$

where $H_{GDBC_{u,v,\alpha,d}}(R_i)$ is the histogram of $GDBC_{u,v,\alpha,d}(z_{i,j})$ in R_i.

In the classification stage, histogram intersection is used to measure the similarity between two histograms $\boldsymbol{H}$ and $\boldsymbol{S}$:

$$P_{HS}(\boldsymbol{H}, \boldsymbol{S}) = \sum_{i=1}^{B} \min(H_i, S_i) \tag{9}$$

where $\boldsymbol{H} = (H_1, H_2, \ldots, H_B)^T$, $\boldsymbol{S} = (S_1, S_2, \ldots, S_B)^T$, B represents the number of histogram bins, and H_i and S_i are the frequency counts in the ith bin.

4. Experiments

The main objectives of the experiments in this section are to evaluate the performance of well-known face recognition algorithms on the PolyU-NIRDF and investigate the strengths and weaknesses of the proposed method and baseline algorithms. We choose the Eigenface (i.e. PCA) (Turk and Pentland, 1991), Fisherface (i.e. LDA) (Belhumeur et al., 1997), LBP (Ahonen et al., 2006), Gabor-PCA, Gabor-LDA and Gabor-LBP as the baseline algorithms. The Gabor-LDA and Gabor-LBP are state-of-the-art methods (Liu and Wechsler, 2002; Zhang et al., 2005) and they are benchmarks to evaluate the performance of FR techniques. Together with the DBC and GDBC, eight methods will be tested on the established PolyU-NIRDF. To further compare DBC and LBP, we conduct another experiment on the visible light band FERET database.

4.1. Experiment 1

In Exp#1, we set up a subset from the whole PolyU-NIRFD database. In this subset, the training set contains 419 frontal images from 138 subjects, while the gallery set and probe set have 574 and 2763 images, respectively. No images in the probe and gallery sets are contained in the training set. The facial portion of each original image is automatically cropped according to the location of the eyes. The cropped face is then normalized to 64×64 pixels. The eight methods are then applied to this dataset to evaluate their performance. For subspace based methods, the distance used in the nearest neighbor classifier is the cosine of the angle between two feature vectors. For LBP and DBC histogram features, we use the histogram intersection similarity measure. The sub-region size and the number of histogram bins for LBP and DBC are 8×8 and 32. Since the Gabor filtering is performed on five scales and eight orientations, we need to downsample the response image to reduce data amount. The downsampling factor is set to 4. Therefore, for Gabor-PCA and Gabor-LDA, the input signal size is $64 \times 64 \times 5 \times 8/4^2 = 10{,}240$.

The FR results by the eight methods are illustrated in Figs. 8 and 9. For subspace based methods PCA, LDA, Gabor-PCA and Gabor-LDA, the curves of recognition rate versus feature vector dimensionality are plotted in Fig. 8. We see that the curves for PCA and LDA are flat when the dimension of feature vector changes from 60 to 120. In this experiment PCA gets similar performance to LDA because the number of training samples for each class is limited. From Figs. 8 and 9 we can clearly see that using Gabor features can improve greatly the performance of all the four methods. For example, the recognition rate of Gabor-LDA is 10% higher than that of LDA. Comparing Fig. 8 with Fig. 9, it can be seen that the LBP and DBC methods achieve better performance than

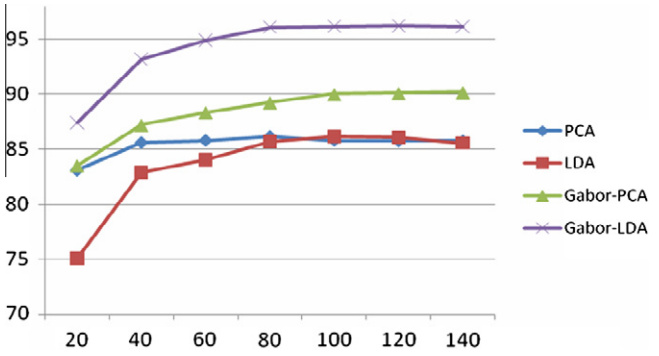

Fig. 8. Recognition results by PCA, LDA, Gabor-PCA and Gabor-LDA in Exp#1.

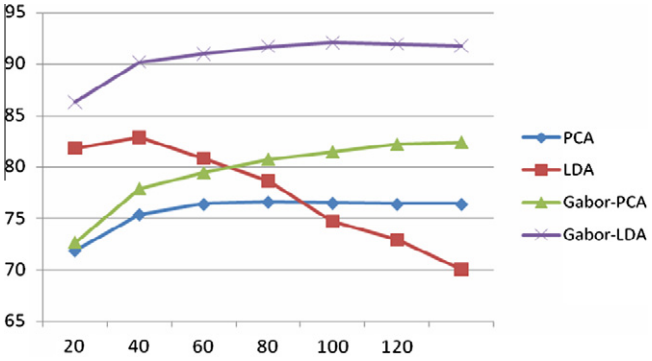

Fig. 10. Recognition results by PCA, LDA, Gabor-PCA and Gabor-LDA in Exp#2.

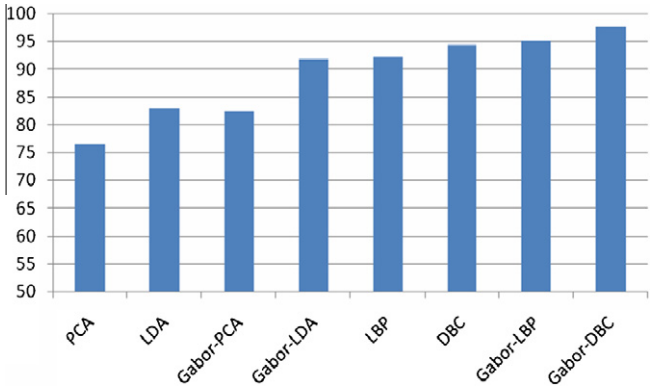

Fig. 9. Recognition results by PCA, LDA, Gabor-PCA, Gabor-LDA, LBP, DBC, Gabor-LBP and Gabor-DBC in Exp#1.

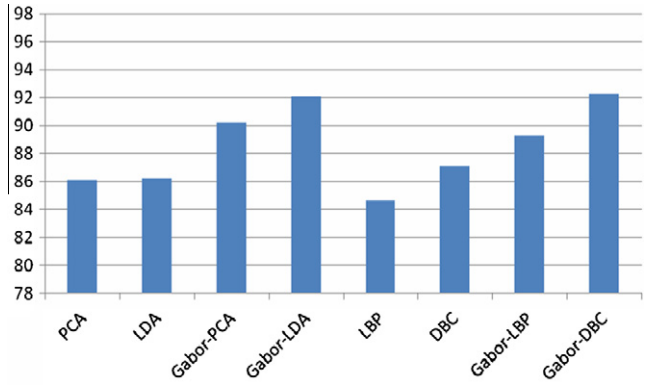

Fig. 11. Recognition results by PCA, LDA, Gabor-PCA, Gabor-LDA, LBP, DBC, Gabor-LBP and Gabor-DBC in Exp#2.

PCA and LDA. Moreover, DBC works better than LBP. This is because DBC can capture the directional edge information and the spatial information in a local region, while LBP only considers the relationship between a given pixel and its surrounding neighbors.

4.2. Experiment 2

In Exp#2, we extracted from the whole database a much bigger subset than in Exp#1. In this subset, the training set contains 1876 frontal images of 150 subjects, while the gallery and probe sets have 1159 and 4747 images, respectively. The recognition results of PCA, LDA, LBP and DBC are illustrated in Figs. 10 and 11. Different from that in Exp#1, LDA achieves much better performance than PCA when using a small number of features. Similar to those in Exp#1, Gabor based methods get much better performances than their original image based counterparts. In this experiment, the Gabor-DBC method achieves an accuracy of 92.3%, while in Exp#1it has an accuracy of 97.6%. This is because the dataset in Exp#2 is more challenging by involving more variations of pose, expression and scale than that in Exp#1.

4.3. Experiment 3

Exp#3 is designed to evaluate the performance of the algorithms on the large variations of pose, expression, illumination, scale, etc. In this subset, training set contains 576 images from 188 subjects, while gallery and probe sets have 951 and 3648 images individually. The results are shown in Figs. 12 and 13. We can see that LBP and DBC have much better performance than PCA and LDA, which validates again that LBP and DBC are effective ways to model NIR face images. DBC performs better than LBP because it exploits the directional information. Particularly, Gabor-DBC achieves the highest recognition rate.

4.4. Experiment 4

To further validate the effectiveness of the proposed DBC scheme, in this sub-section we conduct experiments on the visible light band FERET face database, which is widely used to evaluate face recognition algorithms. The LBP based method achieves state-of-the-art performance on the FERET database (Tan and Triggs, 2007). From the previous experiments we have seen that the LBP methods are more effective than PCA and LDA, thus we only compare DBC with LBP here. In the experiments, the facial portion of each original image is cropped based on the locations of the eyes. The cropped face is then normalized to 88×88 pixels. We

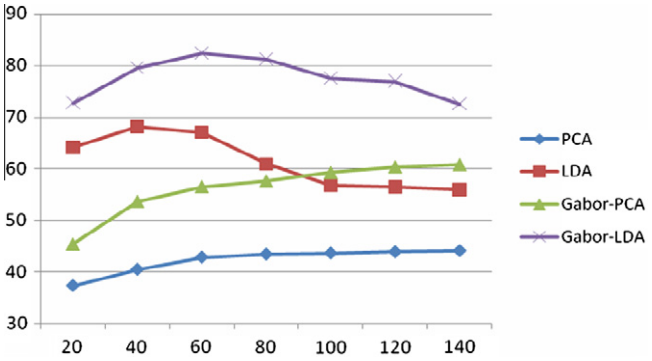

Fig. 12. Recognition results by PCA, LDA, Gabor-PCA and Gabor-LDA in Exp#3.

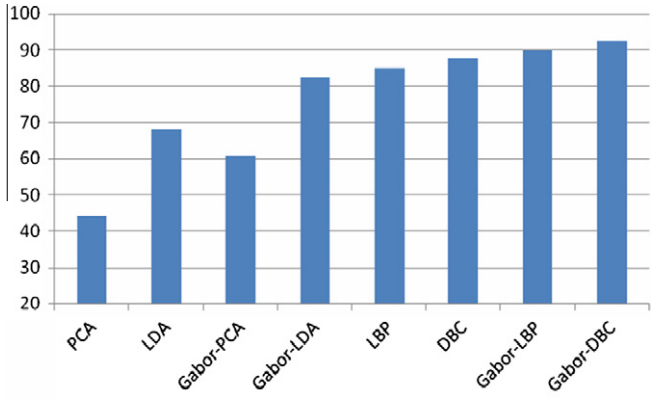

Fig. 13. Recognition results by PCA, LDA, Gabor-PCA, Gabor-LDA, LBP, DBC, Gabor-LBP and Gabor-DBC in Exp#3.

use the same gallery and probe sets as specified in the FERET evaluation protocol. *Fa* that contains 1196 frontal images of 1196 subjects is used as the gallery set, while *Fb* (1195 images with expression variations), *Fc* (194 images taken under different illumination conditions), *Dup I* (722 images taken later in time between one minute to 1031 days) and *Dup II* (234 images, a subset of Dup I taken at least after 18 months) are used as the probe sets.

The experimental results are shown in Figs. 14 and 15. As in the PolyU-NIRFD, we can see that DBC achieves better performance than LBP on the large-scale FERET database. The reason is that DCP can capture the spatial relationship in a local region for a given direction, while LBP only considers the relationship between a given pixel and its surrounding ones.

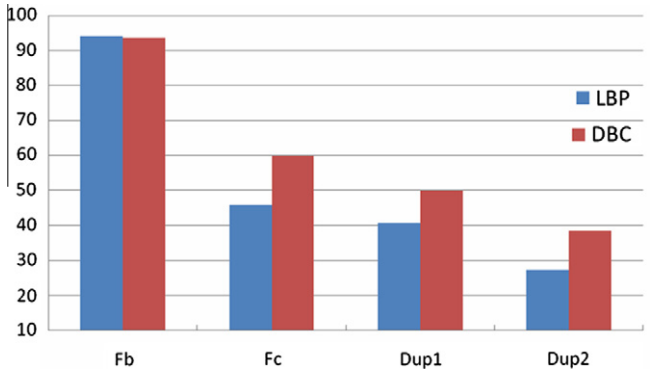

Fig. 14. Recognition results of image based LBP and DBC on FERET.

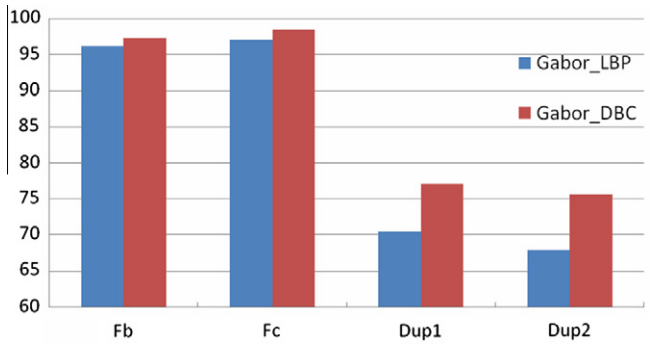

Fig. 15. Recognition results of Gabor based LBP and DBC on FERET.

5. Conclusions

We introduced in this paper the establishment of PolyU near-infrared face database (PolyU-NIRFD), which is one of the largest near-infrared (NIR) face databases so far. A novel NIR-based face recognition method, namely directional binary code (DBC), was also proposed to capture more efficiently the directional edge information in NIR face images. The main characteristics of the PolyU-NIRFD lie in two aspects. First is its large scale. It consists of 35,000 images from 350 subjects. Second is its diversity of variations. It includes variations of pose, expression, illumination, scale, blurring and the combination of them. Comparative study of baseline algorithms and the proposed DBC was performed on the PolyU-NIRFD to verify its effectiveness. Experiments were also performed on the visible light band FERET database to verify the effectiveness of DBC scheme. In the future we will acquire a larger database under both controlled and uncontrolled environment, and investigate more effective NIR face recognition algorithms.

6. Obtaining the PolyU-NIRFD

The information about how to obtain a copy of the database will be found on the website (http://www.comp.polyu.edu.hk/~biometrics/polyudb_face.htm).

Acknowledgments

This work was supported by the GRF grant of HKSAR, the central fund from Hong Kong Polytechnic University, the Natural Science Foundation of China (NSFC) under Contract Nos. 60620160097 and 60903065, the National High-Tech Research and Development Plan of China (863) under Contract Nos. 2006AA01Z193 and 2007AA01Z195, and the Ph.D. Programs Foundation of Ministry of Education of China (No. 20091102120001). Thanks go to Ms. Yafei Chen and Fangyu Li on the experiments.

References

Ahonen, T., Hadid, A., Pietikäinen, M., 2006. Face description with local binary patterns: Application to face recognition. IEEE Trans. Pattern Anal. Machine Intell. 28 (12), 2037–2041.

Enrique, Bailly-Bailliére, Samy, Bengio, Frédéric, Bimbot, Miroslav, Hamouz, Josef, Kittler, Johnny, Mariéthoz, Jiri, Matas, Kieron, Messer, Vlad, Popovici, Fabienne, Porée, Belen, Ruiz, Thiran, Jean-Philippe, 2003. The banca database and evaluation protocol. In: Internat. Conf. on Audio- and Video-based Person Authentication.

Belhumeur, P., Hespanha, J., Kriegman, D., 1997. Eigenfaces vs Fisherfaces: Recognition using class specific linear projection. IEEE Trans. Pattern Anal. Machine Intell. 19, 711–720.

Chellappa, R., Wilson, C.L., Sirohey, S., 1995. Human and machine recognition of faces: A survey. Proc. IEEE 83 (5), 705–740.

Dowdall, J., Pavlidis, I., Bebis, G., 2003. Face detection in the near-ir spectrum. Image Vision Comput. 21, 565–578.

Gao, Y., Leung, M.K.H., 2002. Face recognition using line edge map. IEEE Trans. Pattern Anal. Machine Intell. 24 (6), 764–779.

Gao, W., Cao, B., Shan, S., Chen, X., Zhou, D., Zhang, X., Zhao, D., 2008. The CAS-PEAL large-scale Chinese face database and baseline evaluation. IEEE Trans. Systems Man Cybernet., Part A, 149–161.

Georghiades, A.S., Belhumeur, P.N., Kriegman, D.J., 2001. From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE Trans. Pattern Anal. Machine Intell. 23 (6), 643–660.

Hong, Lin, Wan, Yifei, Jain, Anil K., 1998. Fingerprint image enhancement: Algorithm and performance evaluation. IEEE Trans. Pattern Anal. Machine Intell. 20 (8), 777–789.

Kanekoa, Shunichi, Muraseb, Ichiro, Igarashia, Satoru, 2002. Robust image registration by increment sign correlation. Pattern Recognit 35, 2223–2234.

Kong, Seong G., Heo, Jingu, Boughorbel, Faysal, Zheng, Yue, Abidi, Besma R., Koschan, Andreas, Mingzhong, Yi, Abidi, Mongi A., 2007. Multiscale fusion of visible and thermal IR images for illumination-invariant face recognition. Internat. J. Comput. Vision 71 (2), 215–233.

Lee, K.C., Ho, J., Kriegman, D.J., 2005. Acquiring linear subspaces for face recognition under variable lighting. IEEE Trans. Pattern Anal. Machine Intell. 27 (5), 684–698.

Li, Stan, Chu, Rufeng, Liao, Shengcai, Zhang, Lun, 2007. Illumination invariant face recognition using near-infrared images. IEEE Trans. Pattern Anal. Machine Intell. 29 (4), 627–639.

Liu, C., Wechsler, H., 2002. Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition. IEEE Trans. Image Process. 11 (4), 467–476.

Martinez, M., Benavente, R., 1998. The AR Face Database. CVC, Barcelona, Spain, Tech. Rep. 24.

Messer, K., Matas, J., Kittler, J., Luettin, J., Maitre, G., 1999. XM2VTSDB: The extended M2VTS database. In: Proc. Internat. Conf. on Audio Video-Based Biometric Person Authentication, pp. 72–77.

Ojala, T., Pietikäinen, M., Mäenpää, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Machine Intell. 24 (7), 971–987.

Pan, Z.H., Healey, G., Prasad, M., Tromberg, B., 2003. Face recognition in hyperspectral images. IEEE Trans. Pattern Anal. Machine Intell. 25 (12), 1552–1560.

Philips, P.J., Flynn, P.J., Scruggs, T., Bowyer, K.W., Chang, Jin, Hoffman, Kevin, Marques, Joe, Min, Jaescik, Worek, Willimam, 2005. Overview of the face recognition grand challenges. In: Proc. ICCV, pp. 947–954.

Phillips, P.J., Moon, H., Rizvi, S.A., Rauss, P.J., 2000. The FERET evaluation methodology for face-recognition algorithms. IEEE Trans. Pattern Anal. Machine Intell. 22 (10), 1090–1104.

Phillips, P.J., Grother, P., Micheals, R.J., Blackburn, D.M., Tabassi, E., Bone, J.M. 2003. Face recognition vendor test 2002: Evaluation report. Nat. Inst. Standards Technol., Gaithersburg, MD, Tech. Rep. NISTIR 6965.

Sim, T., Baker, S., Bsat, M., 2003. The CMU pose, illumination, and expression database. IEEE Trans. Pattern Anal. Machine Intell. 25 (12), 1615–1618.

Tan, Xiaoyang, Triggs, Bill, 2007. Fusing Gabor and LBP feature sets for kernel-based face recognition. In: Proceedings of the 3rd International Conference on Analysis and Modeling of Faces and Gestures.

Turk, M.A., Pentland, A.P., 1991. Face recognition using Eigenfaces. In; Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp. 586–591.

Wiskott, L., Fellous, J.M., Krüger, N., von der Malsburg, C., 1997. Face recognition by elastic bunch graph matching. IEEE Trans. Pattern Anal. Machine Intell. 19 (7), 775–779.

Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H. 2005. Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A Novel Non-Statistical Model for Face Representation and Recognition. IEEE International Conference on Computer Vision, 1, 786-791.

Zhao, W., Chellappa, R., Rosenfeld, A., 2003. Face recognition: A literature survey. ACM Comput. Surv. 35, 399–458.

Zou, X., Kittler, J., Messer, K. 2005. Face Recognition Using Active Near-IR Illumination. Proc. British Machine Vision Conf.

Zou, X., Kittler, J., Messer, K., 2007. Illumination invariant face recognition: A survey. In: The First IEEE Internat. Conf. on Biometrics: Theory, Applications, and Systems, pp. 1–8.