Fusing Local Patterns of Gabor Magnitude and Phase for Face Recognition

Shufu Xie, Shiguang Shan, *Member, IEEE*, Xilin Chen, *Senior Member, IEEE*, and Jie Chen, *Member, IEEE*

***Abstract*—Gabor features have been known to be effective for face recognition. However, only a few approaches utilize phase feature and they usually perform worse than those using magnitude feature. To investigate the potential of Gabor phase and its fusion with magnitude for face recognition, in this paper, we first propose local Gabor XOR patterns (LGXP), which encodes the Gabor phase by using the local XOR pattern (LXP) operator. Then, we introduce block-based Fisher's linear discriminant (BFLD) to reduce the dimensionality of the proposed descriptor and at the same time enhance its discriminative power. Finally, by using BFLD, we fuse local patterns of Gabor magnitude and phase for face recognition. We evaluate our approach on FERET and FRGC 2.0 databases. In particular, we perform comparative experimental studies of different local Gabor patterns. We also make a detailed comparison of their combinations with BFLD, as well as the fusion of different descriptors by using BFLD. Extensive experimental results verify the effectiveness of our LGXP descriptor and also show that our fusion approach outperforms most of the state-of-the-art approaches.**

***Index Terms*—Face representation, Fisher's linear discriminant (FLD), fusion, histogram, local Gabor XOR patterns (LGXP).**

I. Introduction

FACE recognition, as one of the most typical applications of image analysis and understanding, has attracted significant attention in many areas such as entertainment, information security, law enforcement, and surveillance [48]. In the last few decades, numerous approaches have been proposed for face recognition and much progress has been made [48]. However, it is still difficult for a machine to recognize human faces accurately under the uncontrolled circumstances. The main challenges include the small interpersonal variations as well as the large intrapersonal variations arising from illumination, pose, expression and other factors.

Manuscript received March 06, 2009; revised September 12, 2009. First published January 26, 2010; current version published April 16, 2010. This work was supported in part by the Natural Science Foundation of China under contracts 60772071, 60833013, and 60872077; in part by the National Basic Research Program of China (973 Program) under contract 2009CB320902; in part by the Hi-Tech Research and Development Program of China under contract 2007AA01Z163; in part by the Grand Program of International S&T Cooperation of Zhejiang Province S&T Department under contract 2008C14063; and in part by the ISVISION Technology Co., Ltd. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Arun Abraham Ross.

S. Xie, S. Shan, and X. Chen are with the Key Laboratory of Intelligent Information Processing of Chinese Academy Sciences (CAS), Institute of Computing Technology, CAS, Beijing, 100190, China (e-mail: sfxie@jdl.ac.cn; sgshan@ict.ac.cn; xlchen@ict.ac.cn).

J. Chen is with the Machine Vision Group, Department of Electrical and Information Engineering, FI-90014, University of Oulu, Finland (e-mail: jiechen@ee.oulu.fi).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2010.2041397

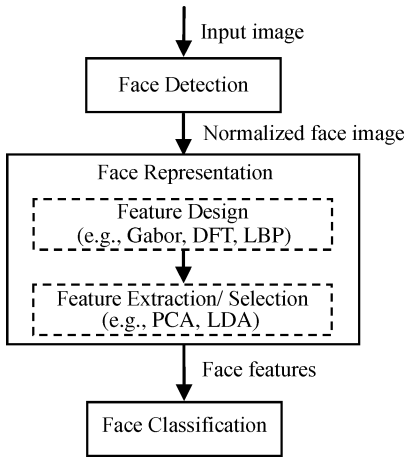


Fig. 1. Architecture of a typical face recognition system. Note: Dashed rectangle represents optional procedure.

Generally speaking, as shown in Fig. 1, a typical face recognition system consists of the following three stages: face detection, face representation and face classification. In this paper, we focus on the second stage, i.e., extracting the internal representation from the normalized face image, which is widely believed the key to the successful face recognition system. As illustrated in Fig. 1, the face representation process might be composed of two sequential procedures: feature design and feature extraction/selection. Specifically, the feature design procedure generally generates a face descriptor from one face image by using some signal processing techniques (e.g., Gabor wavelets [13], discrete Fourier transform (DFT) [17]) or designed operators (e.g., local binary patterns (LBP) [1]). In the second procedure, feature extraction methods commonly generate low-dimensional features by applying subspace analysis (e.g., principal component analysis (PCA) [35] and linear discriminant analysis (LDA) [2]) to the normalized face image or the designed features, while feature selection methods generally select a subset from the input feature set according to some criterion.

Under the above definition, the feature design procedure usually does not depend on any training set or learning process, and the features are commonly empirically designed or derived directly from some signal processing tools. In the earlier years, geometrical features (e.g., nose width and length, mouth position and chin shape) and image template were widely exploited for face recognition. Though intuitively reasonable, they are easily affected by the variations of facial appearances [3]. Later on, transformation features in the frequency domain or wavelet domain attracted much attention for face recognition, such as DFT [17], discrete cosine transform (DCT) [9], discrete wavelet transform (DWT) [5], and Gabor wavelet [6], [13]. Among them, face representation based on Gabor wavelet has been well

known as one of the most successful methods (e.g., [16], [19], [29], and [37]). Recently, motivated by the success of LBP [1], a few methods based on local patterns of Gabor feature have been proposed, such as local Gabor binary patterns (LGBP) [46], [47] and histogram of Gabor phase patterns (HGPP) [43], which are called local Gabor patterns in this paper.

Compared with the feature design procedure, feature extraction procedure generally needs a training set to learn how to distinguish different persons. Among this category, Eigenfaces [35], Fisherfaces [2], and the Bayesian approach [22] are the representative studies, and some variants of them have also been proposed (e.g., [33], [36], and [40]). Essentially, most of these methods belong to either supervised or unsupervised *dimensionality reduction* approaches. Considering the essential nonlinearity of face images, many manifold-related methods have been proposed for feature extraction, such as Isomap [34], locally linear embedding (LLE) [26], locality preserving projections (LPP) [10] and some variants of it (e.g., [4] and [41]). Recently, to unify these dimensionality reduction methods, graph embedding has been proposed as a general framework [39].

Since most feature extraction/selection methods have no special requirement on the input feature, it is reasonable to combine it with feature design procedure in order to seek more effective face representation (e.g., [11], [14], [19], [27], [28], [38], and [42]). Recently, a few methods fusing diverse features have received much attention, such as different frequency bands [12], [21], features in multiple scales [18], global and local features [30], and fusion of Gabor and LBP [32]. Considering that the two parts of Gabor feature (i.e., magnitude and phase) are complementary for face recognition, some approaches based on their fusion have been proposed, such as combining the weak classifiers constructed respectively based on magnitude and phase features [45], the weighted Gabor complex features [8] and the enhanced local Gabor binary patterns [47]. In these works, however, we notice that the phase-based approaches usually achieve worse results than those using magnitude (e.g., [8] and [45]). The reason might be the sensitivity of phase to the varying positions, which leads to severe problems when matching two face images with slight misalignment [16], [37].

To further study the potentials of Gabor phase as well as its fusion with Gabor magnitude, this paper firstly propose local Gabor XOR patterns (LGXP), which encodes Gabor phase by using local XOR pattern (LXP) operator. Then, to reduce the high dimensionality of LGXP descriptor, we propose block-based Fisher's linear discriminant (BFLD) to extract the discriminative low-dimensional features. The BFLD method is borrowed from the previous work in [28], [30], which divides the entire feature set into many feature segments and applies FLD to each segment. Finally, by using BFLD, we fuse local patterns of Gabor magnitude and phase to utilize their complementary information for face recognition. We conduct comparative experimental studies of different local Gabor patterns, their combinations with BFLD as well as their fusion on FERET [24] and FRGC 2.0 [25] databases. Experimental results show that the proposed phase-based method achieves better results than magnitude-based methods in many cases and the introduction of BFLD greatly improves the performances. In addition, on FRGC 2.0 database, the proposed fusion approach achieves results comparable to the best known ones.

The main contribution of this paper lies in two aspects: the LGXP descriptor for Gabor phase encoding and its fusion with local patterns of Gabor magnitude by BFLD. Specifically, the LGXP descriptor, unlike previous studies (e.g., LGBP [46], [47], HGPP [43]), is defined on Gabor phase part with XOR-based local operator. The detailed comparisons with previous work on local Gabor patterns are presented in Section III. Additionally, also different from previous fusion methods (e.g., [8], [45], and [47]), we fuse the local patterns of Gabor magnitude and phase by introducing BFLD to derive the discriminative low-dimensional features.

The rest of this paper is organized as follows. Section II briefly describes the related work. Section III presents the LGXP descriptor. Section IV details our feature extraction approach and Section V presents our fusion approach. Experimental results are presented in Section VI and Section VII gives some conclusions.

II. Related Work

In this section, we firstly describe Gabor wavelet representation and LBP respectively in Sections II-A and II-B. Then, we briefly review the related work about local Gabor patterns in Section II-C.

A. Gabor Wavelet Representation

The Gabor wavelet representation of an image is defined as the convolution of the image with Gabor kernels, i.e.,

$$G_{\mu,\nu}(z) = I(z) * \psi_{\mu,v}(z). \tag{1}$$

Here, $I(\bullet)$ denotes the input image, and $*$ denotes the convolution operator; z denotes the pixel, i.e., $z = (x, y)$, and $\psi_{\mu,v}(\bullet)$ denotes the Gabor kernel with orientation μ and scale ν, which is defined as follows:

$$\psi_{\mu,\nu}(z) = \frac{\|k_{\mu,\nu}\|^2}{\sigma^2} e^{\left(-\|k_{\mu,\nu}\|^2\|z\|^2/2\sigma^2\right)} \left[e^{ik_{\mu,\nu}z} - e^{-\sigma^2/2}\right] \tag{2}$$

where $\|\bullet\|$ denotes the norm operator, and the wave vector $k_{\mu,\nu}$ is defined as follows:

$$k_{\mu,\nu} = k_\nu e^{i\phi_\mu} \tag{3}$$

where $k_\nu = k_{\max}/f^\nu$ and $\phi_\mu = \pi\mu/8$; $k_{\max}$ is the maximum frequency, and f is the spacing between kernels in the frequency domain [19], [37].

For each Gabor kernel, at every image pixel z, a complex number containing two Gabor parts, i.e., real part $\mathrm{Re}_{\mu,\nu}(z)$ and imaginary part $\mathrm{Im}_{\mu,\nu}(z)$, can be generated. Based on these two parts, magnitude $A_{\mu,\nu}(z)$ and phase $\Phi_{\mu,\nu}(z)$ can be computed by (4) and (5), respectively

$$A_{\mu,\nu}(z) = \sqrt{\mathrm{Im}^2_{\mu,\nu}(z) + \mathrm{Re}^2_{\mu,\nu}(z)} \tag{4}$$

$$\Phi_{\mu,\nu}(z) = \arctan\left(\mathrm{Im}_{\mu,\nu}(z)/\mathrm{Re}_{\mu,\nu}(z)\right). \tag{5}$$

B. Local Binary Patterns (LBP)

The LBP operator assigns a label to every pixel of an image by thresholding the 3×3 neighborhood of each pixel with the center pixel value and considering the result as a binary number [23]. For example, as shown in Fig. 2, "11010011" is the designed pattern of the central pixel. By applying LBP operator

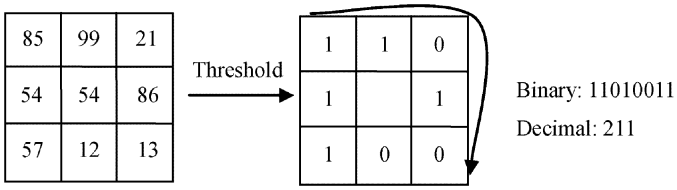


Fig. 2. LBP operator defined in 3 × 3 neighborhood.

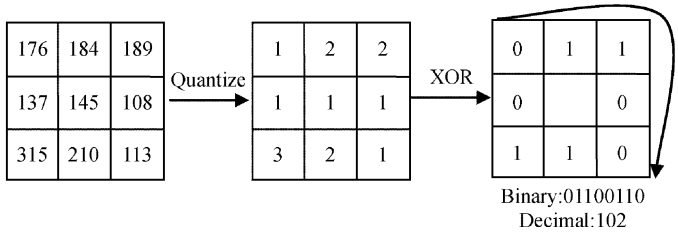


Fig. 3. Example of the encoding method of LGXP where the phase is quantized into 4 ranges.

to one facial image, one pattern map can be computed. Then, the pattern map is divided into many blocks and the histogram computed in each block is concatenated together to form the description of the input facial image [1].

C. Local Gabor Patterns

Recently, a few methods based on local Gabor patterns have been proposed. By encoding local patterns of Gabor magnitude and phase feature via LBP operator, LGBP_Mag [46] and LGBP_Pha [47] have been proposed, respectively. Motivated by the successful applications of Gabor feature in iris recognition [7] and palmprint identification [44], Zhang *et al.* propose local Gabor phase patterns (LGPP) to encode the variations in the local neighborhood of Gabor real and imaginary part by using the LXP operator, which are respectively denoted Re_LGPP and Im_LGPP [43]. Since Gabor wavelet captures salient visual properties such as spatial localization and orientation selectivity, this approach generally performs better than LBP [43].

III. LGXP: Local Gabor XOR Patterns

The basic idea of our method is that, to alleviate the sensitivity of Gabor phase to the varying positions, whether two phases reflect similar local feature should be determined in a "looser" way. Specifically, if two phases belong to the same interval (e.g., $[0°, 90°)$), they are believed similar local features; otherwise, they reflect different local features. In this section, we first present the LGXP descriptor, and then make one comparison with previous local Gabor patterns.

A. LGXP Descriptor

Briefly speaking, as shown in Fig. 3, in our LGXP, phases are firstly quantized into different range, then LXP operator is applied to the quantized phases of the central pixel and each of its neighbors, and finally the resulting binary labels are concatenated together as the local pattern of the central pixel.

Formally, the pattern of LGXP in binary and decimal form is defined as follows:

$$\begin{aligned}\mathrm{LGXP}_{\mu,\nu}(z_c) &= \left[\mathrm{LGXP}^{P}_{\mu,\nu}, \mathrm{LGXP}^{P-1}_{\mu,\nu}, \ldots, \mathrm{LGXP}^{1}_{\mu,\nu}\right]_{\text{binary}} \\ &= \left[\sum_{i=1}^{P} 2^{i-1} \cdot \mathrm{LGXP}^{i}_{\mu,\nu}\right]_{\text{decimal}} \end{aligned} \tag{6}$$

where z_c denotes the central pixel position in the Gabor phase map with scale ν and orientation μ, P is the size of neighborhood, and $\mathrm{LGXP}^{i}_{\mu,\nu}$ $(i = 1, 2, \ldots, P)$ denotes the pattern calculated between z_c and its neighbor z_i, which is computed as follows:

$$\mathrm{LGXP}^{i}_{\mu,\nu} = q\left(\Phi_{\mu,\nu}(z_c)\right) \otimes q\left(\Phi_{\mu,\nu}(z_i)\right), \quad i = 1, 2, \ldots, P \tag{7}$$

where $\Phi_{\mu,v}(\bullet)$ denotes the phase, $\otimes$ denotes the LXP operator, which is based on XOR operator, as defined in (8); $q(\bullet)$ denotes the quantization operator, which calculates the quantized code of phase according to the number of phase ranges, as defined in (9)

$$c \otimes d = \begin{cases} 0, & \text{if } c = d \\ 1, & else \end{cases} \tag{8}$$

$$q\left(\Phi_{\mu,\nu}(\bullet)\right) = i; \quad \text{if } \frac{360 * i}{b} \le \Phi_{\mu,\nu}(\bullet) < \frac{360 * (i+1)}{b},\ i = 0, 1, \ldots, b-1 \tag{9}$$

where b denotes the number of phase ranges.

With the pattern defined above, one pattern map is calculated for each Gabor kernel. Then, each pattern map is divided into m nonoverlapping sub-blocks, and the histograms of all these sub-blocks of all the scales and orientations are concatenated to form the proposed LGXP descriptor of the input face image

$$\begin{aligned} H = [&H_{\mu_0,\nu_0,1}, \ldots, H_{\mu_0,\nu_0,m}; \ldots; \\ &H_{\mu_{o-1},\nu_{s-1},1}, \ldots, H_{\mu_{o-1},\nu_{s-1},m}] \end{aligned} \tag{10}$$

where $H_{\mu,\nu,i}$ $(i = 1, 2, \ldots, m)$ denotes the histogram of the i^{th} sub-block of LGXP map with scale ν and orientation μ. In this paper, Gabor filters of five scales and eight orientations are used. Then, for face recognition, the similarity between two LGXP descriptors H^1 and H^2 can be calculated as follows:

$$S(H^1, H^2) = \sum_{\mu=\mu_0}^{\mu_{o-1}} \sum_{\nu=\nu_0}^{\nu_{s-1}} \sum_{i=1}^{m} \cap\left(H^1_{\mu,\nu,i}, H^2_{\mu,\nu,i}\right) \tag{11}$$

where $\cap$ denotes the histogram intersection operator defined as follows:

$$\cap(h^1, h^2) = \sum_{i=1}^{L} \min\left(h^1_i, h^2_i\right) \tag{12}$$

where L denotes the number of histogram bins.

Obviously, it is important for LGXP to set appropriate b in (9). In order to make the patterns robust to the variations of Gabor phase, the value of b can not be too large. In this study, we find that LGXP performs well enough when $b = 4$. The reason might be that this setting achieves a good balance between the

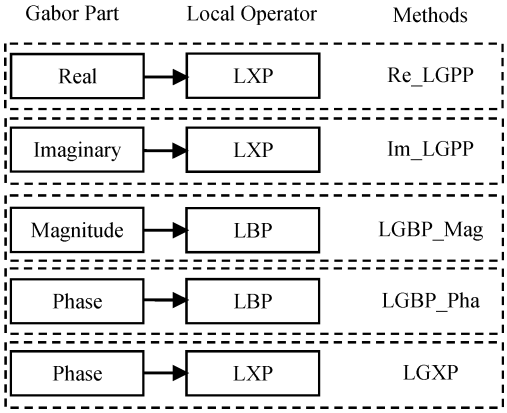


Fig. 4. Comparison of different local Gabor patterns.

robustness to phase variations and representation power of local patterns.

B. Comparisons With Other Local Gabor Patterns

To highlight the difference of the proposed LGXP from previous methods based on local Gabor patterns, we illustrate their brief flowcharts in Fig. 4. As shown in the figure, these methods differ in either the Gabor part or the local operator. Specifically, Re_LGPP and Im_LGPP are respectively based on the real and imaginary parts followed by LXP operator [43]; LGBP_Mag [46] and LGBP_Pha [47] use the LBP operator to encode respectively the magnitude and phase part, while the proposed LGXP exploits the phase part and the LXP operator.

Compared with LGBP_Pha, attributing to the large-step quantizing operation in our method, the LGXP is expected to be more robust to the phase variations due to the varying position. Even when a phase changes significantly, its quantized code keeps invariant if only it is still within the same range. However, for LBP, the pattern is easily affected when the order relationship of two phases changes, which occurs frequently especially in the near-uniform regions [31].

In some sense, LGXP can be seen as a fusion of Re_LGPP and Im_LGPP at the "encoding" level. Intuitively, the four phase angle ranges (i.e., $[0°, 90°), [90°, 180°), [180°, 270°), [270°, 360°)$) correspond to the four combination of the signs of real and imaginary parts in the LGPP method (i.e., "00", "10", "11", "01", where "0" and "1" denote the signs of real and imaginary parts, respectively). As a combination, the patterns in LGXP reflect more local variations than those in Re_LGPP and Im_LGPP.

IV. Feature Extraction Using Block-Based FLD

As one feature design approach, the proposed LGXP descriptor can be directly applied to face recognition by using (11) as similarity measurement. However, this is not good enough since its feature dimension (i.e., $\mathrm{m} \times 40 \times 2^{\mathrm{P}}$) is very high due to the use of multiple Gabor filters (e.g., 40 in this study). In theory, to reduce the dimensionality, we can apply FLD directly; whereas, in the case of so high-dimensional feature, FLD also suffers from heavy "small sample size (SSS)" problem. Therefore, we further present the block-based FLD (BFLD) approach.

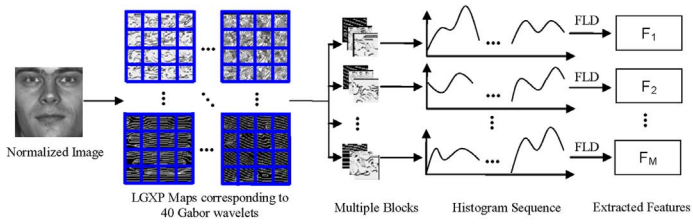


Fig. 5. Flowchart of the BFLD feature extraction approach.

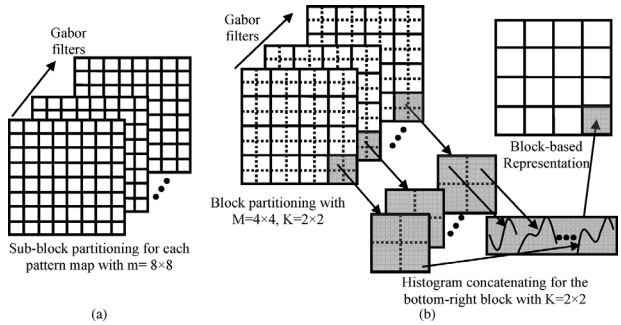


Fig. 6. Block partition strategy in our approaches. (a) LGXP. (b) BFLD.

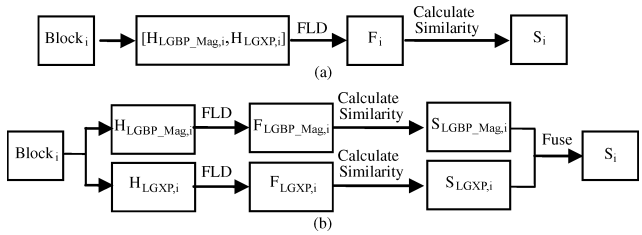


Fig. 7. Fusion of LGBP_Mag and LGXP. (a) Feature-level fusion. (b) Score-level fusion.

A. Block-Based FLD

The basic idea of BFLD is firstly to divide the high-dimensional LGXP descriptor into multiple feature segments (corresponding to different spatial blocks in the face image), then apply FLD to each segment, and finally combine the decisions of all the block-wise FLD. By such a "*divide and conquer*" strategy, the SSS problem is greatly weakened since the dimensionality of the input feature for each FLD is much lower.

Fig. 5 illustrates the flowchart of the BFLD approach. Briefly speaking, for each face image, we firstly calculate its multiple LGXP maps. Then, we divide these pattern maps into multiple nonoverlapping blocks and calculate the block based representations. Based on the training set, we learn the FLD matrices to calculate the low-dimensional features for each block. For clarity, we describe the FLD matrices learning procedure and feature extraction procedure in Algorithms 1 and 2, respectively.

Algorithm 1 Procedure of FLD matrices learning

Input: T: the training set with N normalized face images; M: number of blocks for FLD learning; K: number of sub-blocks per block for histogram extraction; P: size of neighborhood for LGXP.

Output: M FLD matrices W_i $(i = 1, 2, \ldots, M)$.

Step 1 For each image $I \in T$, compute its block-based representation according to Steps 1.1 $\sim$ 1.3:
Step 1.1 Compute its LGXP maps $\Re_{\mu,\nu}$ $(\mu = \mu_0, \mu_1, \ldots, \mu_{o-1}, \nu = \nu_0, \nu_1, \ldots, \nu_{s-1})$.
Step 1.2 Divide $\Re_{\mu,\nu}$ into M nonoverlapping blocks $\Re_{\mu,\nu,i}$ $(i = 1, 2, \ldots, M)$, and further partition each block $\Re_{\mu,\nu,i}$ into K nonoverlapping sub-blocks $\Re^j_{\mu,\nu,i}$ $(j = 1, 2, \ldots, K)$. For sub-block $\Re^j_{\mu,\nu,i}$, calculate its histogram $H^j_{\mu,\nu,i}$.
Step 1.3 For each block $\Re_i$ $(i = 1, 2, \ldots, M)$, concatenate the histograms of all its sub-blocks at all the scales and orientations to get one histogram sequence: $H_i = [H_{\mu_0,\nu_0,i}, \ldots, H_{\mu_0,\nu_{s-1},i}; \ldots; H_{\mu_{o-1},\nu_0,i}, \ldots, H_{\mu_{o-1},\nu_{s-1},i}]$, where $H_{\mu,\nu,i}$ denotes the concatenated histograms of its K sub-blocks, i.e., $H_{\mu,\nu,i} = [H^1_{\mu,\nu,i}, H^2_{\mu,\nu,i}, \ldots, H^K_{\mu,\nu,i}]$. Thus, image I is represented as M histogram sequences H_i $(i = 1, 2, \ldots, M)$.
Step 2 Learn FLD matrices for each block by using training set T according to Steps 2.1 $\sim$ 2.2:
Step 2.1 Obtain M feature sets S_i $(i = 1, 2, \ldots M)$ by collecting the features of the same spatial block from each of the N training images, i.e., $S_i = \{H_{1,i}, H_{2,i}, \ldots, H_{N,i}\}$, where $H_{j,i}$ $(j = 1, 2, \ldots, N)$ denotes the histogram feature from the i^{th} block of the j^{th} training image.
Step 2.2 Based on each S_i (i.e., the training set for the ith block), learn an FLD matrix W_i according to the mode in Fisherfaces [2], i.e., the histogram vectors are first projected to a PCA subspace and the transformed features are then used to learn the matrix.

Algorithm 2 Feature extraction using FLD matrices

Input: Normalized face image I; M, K and P, with the same meanings as those in Algorithm 1; the learned FLD matrices W_i $(i = 1, 2, \ldots, M)$.

Output: M low-dimensional vectors F_i $(i = 1, 2, \ldots, M)$.

Step 1 For face image I, calculate its block-based representation H_i $(i = 1, 2, \ldots, M)$ according to Steps 1.1 $\sim$ 1.3 in Algorithm 1.
Step 2 Calculate its M low-dimensional vectors F_i using the linear transforms: $F_i = (W_i)^T H_i$.

B. Block Partition Strategy

This part describes the block partition strategy used in LGXP and BFLD, which is respectively illustrated in Fig. 6(a) and (b). Specifically, in LGXP, each pattern map is directly divided into m sub-blocks; while in BFLD, each pattern map is divided into M blocks and each block is further partitioned into K sub-blocks. For BFLD, this strategy preserves the spatial relations of different sub-blocks, and obtains better representation than using only one histogram. In some sense, one block in BFLD can be viewed as the combination of some sub-blocks in LGXP, and they satisfy the following equation when the appropriate values are set for them: $m = M \times K$. To make the sub-block partition clear, we adopt "$a \times b$" to denote the number of sub-blocks, where a is the number of sub-blocks in horizontal direction and b that in vertical direction. The similar notations are also used in BFLD, as shown in Fig. 6.

V. Fusing Local Patterns of Gabor Magnitude and Phase for Face Recognition

Gabor magnitude and phase provide complementary information to distinguish different human faces. Therefore, to further improve the recognition accuracy, this section presents how to combine the local patterns of Gabor magnitude and phase, i.e., LGBP_Mag and LGXP. Briefly speaking, we fuse them by using BFLD at two levels, i.e., feature-level and score-level, as illustrated in Fig. 7.

Both feature-level and score-level fusion methods are conducted on block-level. As shown in Fig. 7(a), in the case of feature-level fusion, for each block, the histogram representations of LGBP_Mag and LGXP are simply concatenated into one vector, which is then used to extract feature by FLD; based on the extracted features, one similarity can be calculated. In contrast, for the score-level fusion, as shown in Fig. 7(b), for each block, two low-dimensional vectors are respectively extracted from the histogram representations of LGBP_Mag and LGXP by FLD, and then used to compute two separate similarity scores; finally, these two scores are fused together as the final score.

Formally, these two fusion approaches are described in details as follows.

1) **Calculating block-based representation**. For each image, calculate its block-based representation of LGBP_Mag and LGXP according to Algorithm 1 Steps 1.1 $\sim$ 1.3: $\mathrm{H}_{\mathrm{LGBP\_Mag,i}}$ and $\mathrm{H}_{\mathrm{LGXP,i}}$ $(i = 1, 2, \ldots, M)$, where M denotes the number of blocks per pattern map.
2) **Fusing LGBP_Mag and LGXP**.
   i) **Feature-level Fusion**. As shown in Fig. 7(a), for the ith block, we represent it as one vector by concatenating its LGBP_Mag histograms and LGXP histograms, i.e., $\mathrm{H_i} = [\mathrm{H}_{\mathrm{LGBP\_Mag,i}}, \mathrm{H}_{\mathrm{LGXP,i}}]$. Then, by using the similar procedures as in Algorithms 1 and 2, we calculate its low-dimensional feature vector F_i. For the gallery block I^g_i and its corresponding probe block I^p_i, the similarity between them is calculated as follows:

$$S(I^g_i, I^p_i) = sim(F^g_i, F^p_i) = \frac{F^g_i \cdot F^p_i}{\|F^g_i\| \cdot \|F^g_i\|} \qquad (13)$$

where F_i^g and F_i^p respectively denote the low-dimensional features of I_i^g and I_i^p, and cosine similarity is adopted.

ii) **Score-level Fusion**. As shown in Fig. 7(b), from the two histogram sequences of the ith block, $\mathrm{H}_{\mathrm{LGBP\_Mag,i}}$ and $\mathrm{H}_{\mathrm{LGXP,i}}$, we respectively extract their low-dimensional FLD features, $F_{\mathrm{LGBP\_Mag,i}}$ and $F_{\mathrm{LGXP,i}}$. Then, we respectively compute two similarities between the gallery block I_i^g and its corresponding probe block I_i^p according to (13), namely, $S_{\mathrm{LGBP\_Mag,i}}$ and $S_{\mathrm{LGXP,i}}$. Finally, these two similarities are fused together according to the weighted sum rule

$$S\left(I_i^g, I_i^p\right) = w \cdot S_{\mathrm{LGBP\_Mag,i}} + (1-w) \cdot S_{\mathrm{LGXP,i}} \quad (14)$$

where w denotes the weight of LGBP_Mag, and varies within [0, 1].

3) **Calculating the similarity between any two face images**. Based on the block-wise similarities computed in Step 2, the similarity between the gallery and probe face images is calculated by fusing the similarities of the corresponding blocks according to the sum rule [15]

$$S(I^g, I^p) = \frac{1}{M}\sum_{i=1}^{M} S\left(I_i^g, I_i^p\right). \quad (15)$$

4) **Face classification**. For identification tasks, the identity of the input unknown face is usually classified as the face which is the most similar to the input one, or recognized as "unknown person" if this maximal similarity is smaller than a given threshold; whereas for verification tasks, the similarity between the query face and the claimed target face is first computed, and then the claimed identity of the query face is confirmed or rejected by judging the similarity larger or smaller than a predefined threshold.

For both fusion methods, in principle, the parameters (e.g., K and P) of LGBP_Mag and LGXP can take different values since they are calculated individually. However, this implies more parameters for us to empirically tune. So, the corresponding parameters of LGBP_Mag and LGXP are set the same, i.e., $K_{\mathrm{LGBP\_Mag}} = \mathrm{K}_{\mathrm{LGXP}}$ and $\mathrm{P}_{\mathrm{LGBP\_Mag}} = \mathrm{P}_{\mathrm{LGXP}}$. In addition, for the score-level fusion, the weighted sum rule is adopted in (14) to account for different contributions of LGBP_Mag and LGXP for classification. Actually, our experimental results show that nearly equal weights (e.g., 0.5) perform well, which also implies that the proposed phase-based LGXP method performs equally well compared with magnitude-based LGBP.

VI. Experiments

In this section, we evaluate our approach on FERET [24] and FRGC 2.0 databases [25]. First, we briefly describe experimental conditions. Then, we evaluate our approach with different parameters using two probe sets of FERET. Finally, we make a detailed comparison between our approach and some state-of-the-art approaches on these two databases.

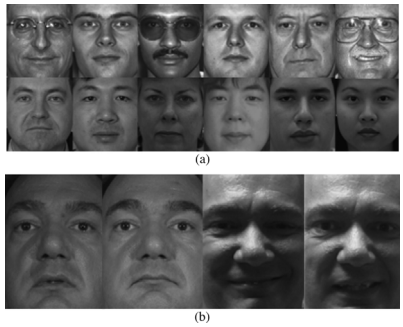


Fig. 8. Some examples of normalized face images in our experiments. (a) FERET. (b) FRGC 2.0.

A. Databases and Experimental Setup

The FERET is a general evaluation designed to measure the performance of laboratory algorithms. Please refer to [24] for details. In our experiments, based on the standard gallery (1,196 images of 1,196 subjects), we test the recognition rates of our approaches on the four probe sets: Fb (1,195 images of 1,195 subjects), Fc (194 images of 194 subjects), Duplicate I (abbreviated as "DupI", 722 images of 243 subjects), and Duplicate II (abbreviated as "DupII", 234 images of 75 subjects). In our experiments, 1,002 frontal images of 429 subjects in the FERET training CD are used as the training set to learn FLD matrices. As shown in Fig. 8(a), all the face images are aligned based on the manually located eye centers provided by the original database, and then normalized to 80 $\times$ 88 pixels. No photometric normalization is conducted for all the methods under comparison, and nearest neighbor classifier is adopted for the final classification. For the Gabor filters, the parameters are set as follows: $\mu \in \{0, 1, \ldots, 7\}$, $\nu \in \{0, 1, \ldots, 4\}$, $\sigma = 2\pi$, $k_{\max} = \pi$, $f = \sqrt{2}$, and the size of the filter window is set to 32 $\times$ 32 pixels.

The Face Recognition Grand Challenge (FRGC) is designed to determine the merits of face recognition techniques by presenting to researchers a six-experiment challenge problem along with data corpus of 50,000 images [25]. For each experiment, three Receiving Operator Characteristic (ROC) curves are generated. ROC1 corresponds to the images collected within semesters, ROC2 within a year, and ROC3 between semesters. In our experiments, we use the given training set (12,776 images of 222 subjects) to learn the FLD matrices and carry out Experiment 1 and 4 for evaluations. Since FRGC intends to evaluate the effectiveness of high-resolution, in our experiments, all the face images are aligned based on the manually located eye centers provided by the original database and normalized to a larger size, i.e., 128 $\times$ 168 pixels. Some examples are shown in Fig. 8(b). Correspondingly, the size of the Gabor filter window is set to 64 $\times$ 64 pixels and $k_{\max} = \pi/2$. In our following experiments, unless otherwise stated, no photometric preprocessing is conducted.

B. Parameter Selection

In this section, by using Fb and DupI probe sets in FERET, we evaluate how the free parameters influence the proposed

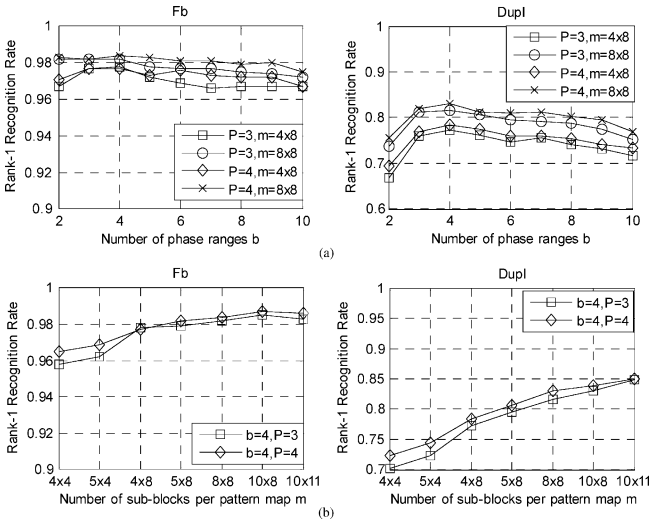


Fig. 9. Effect of different parameters of LGXP approach when tested on FERET probe sets.

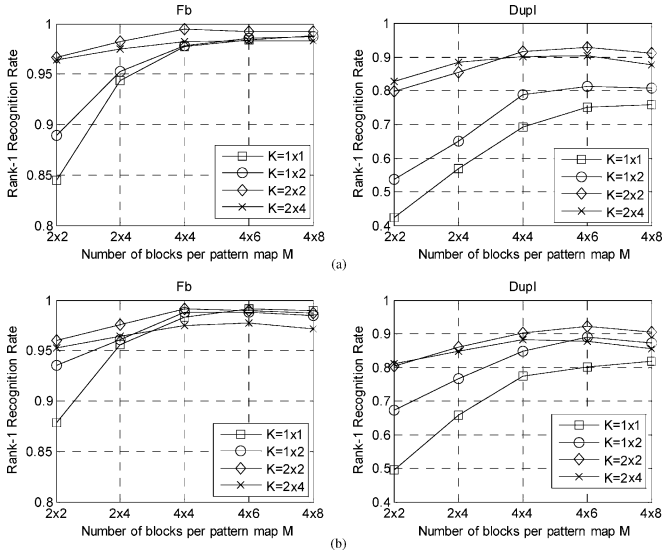


Fig. 10. Effects of different parameters of "LGXP+BFLD" when tested on FERET probe sets. (a) $\mathrm{P} = 3$. (b) $\mathrm{P} = 4$.

method, namely, LGXP, "LGXP+BFLD", the feature-level fusion method (denoted as "F[LGBP_Mag+LGXP]") and score-level fusion method (denoted as "S[LGBP_Mag+LGXP]").

1) Performances of LGXP With Different Parameters: For LGXP, it has three free parameters, i.e., number of phase ranges b, number of sub-blocks per pattern map m and size of neighborhood P. We analyze their meanings as follows: the parameter b determines whether two phases are similar enough, so, it forms a tradeoff between robustness and accuracy; the parameter m affects the relationship of different facial parts as well as the final feature dimensionality, and it can balance between the spatial information preserving and the feature dimension; the parameter P decides how much local variation is encoded in LGXP, and a larger P leads to more elaborate local patterns. However, this also implies higher feature dimension, since the number of the histogram bins is 2^P and the final dimension of the LGXP feature is $40 \times m \times 2^P$.

Based on the above analysis, we empirically set the value of b as one integer within [2, 10], set the minimal size of sub-block to 8×8 pixels to make the sub-block large enough for histogram estimation, and set the value of P to be 3 and 4, which respectively results in 8 or 16 kinds of pattern totally. Our experience also indicates that how to select these 3 or 4 neighbors brings no much difference in performances, so we select them evenly from 3×3 neighborhood.

So, we sample some parameter combinations and conduct experiments to show their performances, as shown in Fig. 9. In Fig. 9(a), we plot how performance varies with b, when $P = 3$ or 4 and $m = 4 \times 8$ or 8×8. From the figure, it can be seen that the best result is achieved when $b = 4$. The reason might be that the patterns encoded by four phase ranges achieve one good tradeoff between robustness to phase variations and representation ability of local patterns. We also notice that, the larger P and m (e.g., $P = 4$, $m = 8 \times 8$) lead to better results than the smaller ones (e.g., $P = 3$, $m = 4 \times 8$), and one limitation is higher dimensional features used. Fig. 9(b) illustrates how the performance of LGXP changes with m when $b = 4$ and $P = 3$ or 4. It is clear that the performances advance roughly with the increase of m. We owe this observation to its balance between spatial information preservation and robustness to the variations in sub-block.

2) Performances of "LGXP+BFLD" With Different Parameters: Except b and P, there are two more parameters for "LGXP+BFLD": number of blocks per pattern map M and number of sub-blocks per block K. In some sense, the product of M and K correspond to m in LGXP (i.e., $m = M \times K$, referring to Fig. 6 in Section IV-B). In addition, K and P together determine the feature dimension (i.e., $40 \times K \times 2^P$) of each block, which can not be very high to avoid the large computational cost for FLD learning. To show how the performance varies with the parameters, we empirically set the parameters as follows: the value of M varies from 2×2 to 4×8, K varies from 1×1 to 2×4, b is fixed to 4, P is set to 3 or 4, and the FLD dimensionality of each block is set to 200.

Fig. 10(a) and (b) illustrates the results of "LGXP+BLFD" on FERET Fb and DupI sets with different parameters. As shown in the figure, the performances roughly increase with the increase of M especially when K is small (e.g., 1×1 and 1×2). We also note that, when $M = 16$ (i.e., 4×4) or 24(i.e., 4×6), $K = 4$ (i.e., 2×2), and $P = 3$, our method achieves peak performance. In this case, each block can be represented effectively by the concatenated histograms of all its sub-blocks. Thus, each block has strong representation power and their fusion achieves better results.

Based on the results in Figs. 9 and 10, we briefly summarize the appropriate setting of the parameters for "LGXP+BFLD" as follows: $b = 4$, $P \in \{3, 4\}$, $K \in \{4, 8\}$, while M should be adapted according to the image size, and our experiences also show that it performs reasonably well when $M = 16$ in most cases.

3) Comparisons Between BFLD-Based Methods and Fusion Approaches: This sub-section mainly evaluates the performance of the proposed fusion methods. Based on the above analysis, we simply set the parameters of our method as follows: $M = 16$ (i.e., 4×4), $K = 4$ (i.e., 2×2), and P is fixed to 3 or 4.

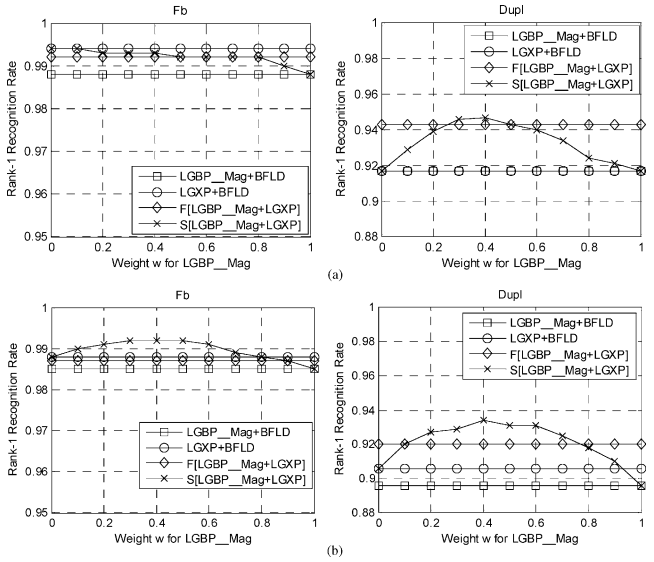


Fig. 11. Rank-1 recognition rates of the proposed BFLD-based approaches on FERET probe sets. (a) $\mathrm{P} = 3$. (b) $\mathrm{P} = 4$.

For our score-level fusion method, i.e., S[LGBP_Mag+LGXP], we evenly sample different weights within [0, 1]. The value of b is set to 4 for LGXP. Fig. 11(a) and (b) illustrates the results of different methods when P is set to 3 and 4 respectively.

From Fig. 11, as expected, both fusion methods outperform nonfusion methods, especially on DupI probe set, and the two fusion methods perform comparably well. As to the weight for LGBP_Mag in the score-level fusion approach, the best result is achieved when its value is around 0.5. Since both "LGBP_Mag+BFLD" and "LGXP+BFLD" approaches achieve comparable results, it is reasonable to fuse their scores by regarding nearly equal weights as their prior probabilities.

C. Evaluation of Proposed Approaches on FERET Database

In this part, we evaluate our methods on the four FERET probe sets, and make a detailed comparison with some state-of-the-art approaches. The parameters for local Gabor patterns are as follows: $m = 64$ (i.e., 8×8), $P = 4$. The parameters for BFLD-based approaches are as follows: $M = 16$ (i.e., 4×4), $K = 4$ (i.e., 2×2), $P = 4$, and the FLD dimensionality of each block is selected when the best result is achieved. The value of b is set to 4 for LGXP, and the weight for score-level fusion approach is set to 0.5. Since score-level fusion approach achieves comparable even better results than the feature-level fusion approach in most cases, we only report the results of the score-level fusion of other descriptors.

Table I tabulates the results of different local Gabor patterns, their combinations with BFLD as well as their fusion approaches on FERET probe sets. In Fig. 12, we plot the cumulative match curves of different methods on FERET probe sets. All the methods are implemented by us, and are conducted on the same image set. From Table I and Fig. 12, we have the following observations: *first*, the proposed phase-based LGXP method outperforms other previous local Gabor pattern methods; *second*, by combining BFLD, the performances of all local Gabor pattern methods are impressively improved; *third*, fusing various Gabor parts in score-level or feature-level can further improve performance, and the score-level fusion of LGBP_Mag and LGXP produces nearly the best results on all the four probes. These observations strongly support the effectiveness of the techniques proposed in this study. Especially, the proposed phase-based method has outperformed those magnitude-based methods, which implies that it is possible to present one effective face descriptor based on Gabor phase if its limitation is alleviated properly.

Except comparing our methods with other local Gabor pattern methods, as shown in Table II, comparisons are also made with other state-of-the-art results reported recently by other researchers on these FERET probe sets. In the table, all the results for comparison are directly cited from papers published recently. From the comparison, we can see that our S[LGBP_Mag+LGXP] method has made impressive improvement especially on the DupI and DupII probe sets, which further validates the effectiveness of the proposed method.

Clearly, compared with other methods based on local Gabor patterns (e.g., LGBP and LGPP), LGXP is almost of equal complexity. However, after combination with BFLD and fusion, computational complexity does increase a lot. Fortunately, the M BFLD matrices are learned off-line, and its application in online stage is of low complexity. Besides, compared with local Gabor pattern method without FLD, the dimensionality of the final feature for each face image is much lower. Thus, computational cost is reduced accordingly when matching against the face database. In sum, compared with the accuracy enhancement, the increase in computational complexity is acceptable.

D. Evaluation of Proposed Approaches on FRGC2 Database

In this section, we conduct a series of experiments on the FRGC 2.0 database to further evaluate different approaches. The parameters for local Gabor pattern methods are as follows: $m = 64$ (i.e., 8×8), $P = 4$, and the parameters for BFLD based approaches are set as follows: $M = 16$ (i.e., 4×4), $K = 8$ (i.e., 4×2), $P = 3$, and the FLD feature dimension of each block is set to be about 200. The value of b is also set to 4 for LGXP, and the weight w is set to 0.5 in the score-level fusion method. Additionally, to avoid the high-dimensionality problem for learning FLD matrices, in our feature-level fusion method, K is set to 4.

Similar to Table I, with the above parameters, we make one comparison between different methods in Table III. Note that, in this experiment, the performance measurement is Verification Rate (VR) at False Accept Rate (FAR) equal to 0.1%, which is quite different from the rank-1 recognition rate on FERET database. Fig. 13 plots the ROC curves of different methods with the same parameter values as in Table III. For clarity, only the curves corresponding to ROC3 are illustrated.

Compared with the results in Table I, Table III further exhibits several interesting observations. *First*, the proposed LGXP method performs the best in Experiment 1, which is a relatively easy test. But, in Experiment 4, its performances are a little worse than those of LGBP_Pha. More important, all the methods based on pure Local Gabor patterns perform surprisingly bad (mostly lower than 20%) in Experiment 4. This might be accounted for by the nondiscriminative nature

TABLE I
RANK-1 RECOGNITION RATES OF DIFFERENT APPROACHES ON FERET DATABASE

Method Category	Methods	FERET Probe Sets			
		Fb	Fc	DupI	DupII
Local Gabor Patterns	LGBP_Mag	96%	96%	74%	70%
	LGBP_Pha	97%	97%	80%	79%
	Re_LGPP	**98%**	99%	80%	78%
	Im_LGPP	**98%**	99%	76%	73%
	LGXP	**98%**	**100%**	**82%**	**83%**
Combinations with BFLD	LGBP_Mag+BFLD	**99%**	99%	91%	**91%**
	LGBP_Pha+BFLD	**99%**	99%	88%	86%
	Re_LGPP+BFLD	**99%**	**100%**	91%	87%
	Im_LGPP+BFLD	**99%**	**100%**	90%	88%
	LGXP+BFLD	**99%**	**100%**	**92%**	**91%**
Fusion of Different Gabor Parts	S[LGBP_Mag+LGBP_Pha]	**99%**	99%	92%	92%
	S[Re_LGPP+Im_LGPP]	**99%**	**100%**	92%	89%
	S[LGBP_Mag+Re_LGPP]	**99%**	99%	93%	91%
	F[LGBP_Mag+LGXP]	**99%**	99%	93%	92%
	S[LGBP_Mag+LGXP]	**99%**	99%	**94%**	**93%**

Note: Bold value denotes the best result in each method category and the similar notations are also used in the following tables.

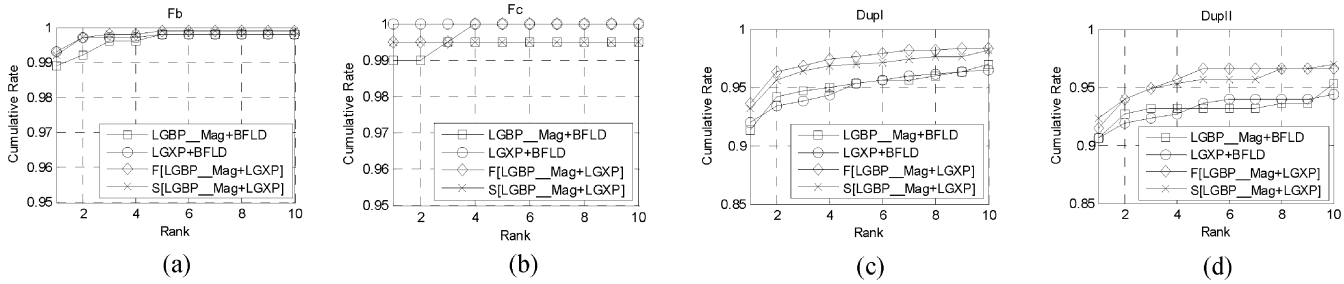


Fig. 12. Cumulative match curves of different methods on FERET probe sets.

TABLE II
COMPARISONS WITH SOME STATE-OF-THE-ART METHODS ON FERET DATABASE

Methods	FERET Probe Sets			
	Fb	Fc	DupI	DupII
Gao et al's Method in [8]	98.0%	83.0%	70.0%	56.0%
Tan and Triggs's Method in [32]	98.0%	98.0%	90.0%	85.0%
Zhang et al's Method in [47]	99.0%	96.0%	78.0%	77.0%
Zou et al's Method in [49]	**99.5%**	**99.5%**	85.0%	79.5%
Proposed score-level fusion method	99.0%	99.0%	**94.0%**	**93.0%**

Note: As is in this work, the eye positions in the original face images are manually located for the methods in [32], [47] and [49], and unknown for that in [8].

of these methods, which can hardly handle challenging variations due to complex illumination and blurring effect in the images of Experiment 4 (i.e., Query set). *Second*, very exciting improvement is achieved from the introduction of BFLD for all local Gabor pattern methods, especially on the Experiment 4 (from 20% to 80%). We attribute this improvement to the discriminative nature and "divide and conquer" methodology of BFLD. Specifically, as a kind of statistical learning method, by modeling the extrinsic variations in different images of the same person as within-class variation and simultaneously modeling the intrinsic variations in identity as between-class variations, FLD can very effectively learn how to distinguish different persons, if only the extrinsic variation in the testing set are similar to those in the training set; in addition, as one "divide and conquer" method, BFLD actually ensembles many component FLD and ensemble integration might also lead to much improvement in performance. *Finally*, fusion of different local Gabor patterns might lead to further improvement, and our score-level fusion of LGBP_Mag and LGXP reports the best results in both Experiment 1 and 4. However, we also notice that fusion of different Gabor parts does not necessarily bring improvement, for instance, S[LGBP_Mag+Re_LGPP] is slightly worse than "LGBP_Mag+BFLD".

Considering that face images in FRGC 2.0 database display complex illumination variations, two photometric normalization methods, Histogram Equalization (abbreviated as "HE") and the preprocessing method in [31] (abbreviated as "PP"), are tested and compared with no photometric normalization in Table IV.

TABLE III
VERIFICATION RATES OF DIFFERENT METHODS ON FRGC 2.0 DATABASE WHEN FAR = 0.1%

Method Category	Methods	Exp.1			Exp.4		
		ROC1	ROC2	ROC3	ROC1	ROC2	ROC3
Local Gabor Patterns	LGBP_Mag	80.3%	76.3%	71.9%	11.3%	11.4%	11.7%
	LGBP_Pha	88.1%	84.5%	80.4%	**20.7%**	**19.5%**	**19.4%**
	Re_LGPP	90.0%	86.4%	83.0%	18.3%	17.9%	17.4%
	Im_LGPP	90.1%	86.5%	83.0%	18.5%	18.0%	17.6%
	LGXP	**90.7%**	**87.6%**	**84.3%**	19.0%	18.4%	17.8%
Combinations with BFLD	LGBP_Mag+BFLD	98.1%	97.5%	96.8%	**79.6%**	**80.3%**	**81.0%**
	LGBP_Pha+BFLD	97.4%	96.5%	95.6%	71.5%	72.0%	72.3%
	Re_LGPP+BFLD	98.2%	97.4%	96.6%	73.8%	74.1%	74.4%
	Im_LGPP+BFLD	98.2%	97.4%	96.7%	74.1%	74.4%	74.7%
	LGXP+BFLD	**98.5%**	**97.9%**	**97.3%**	78.2%	78.6%	78.9%
Fusion of Different Gabor Parts	S[LGBP_Mag+LGBP_Pha]	98.3%	97.6%	97.0%	81.5%	82.3%	83.1%
	S[Re_LGPP+Im_LGPP]	98.4%	97.7%	97.0%	77.2%	77.7%	78.0%
	S[LGBP_Mag+Re_LGPP]	97.7%	97.1%	96.4%	79.5%	79.4%	79.5%
	F[LGBP_Mag+LGXP]	98.6%	98.0%	97.3%	81.9%	82.6%	83.1%
	S[LGBP_Mag+LGXP]	**98.7%**	**98.1%**	**97.5%**	**83.6%**	**84.3%**	**84.9%**

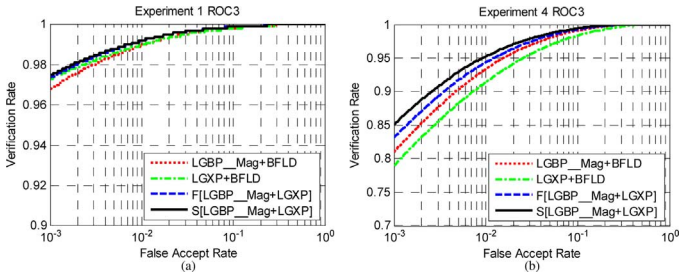


Fig. 13. ROC curves of different methods on FRGC 2.0 database.

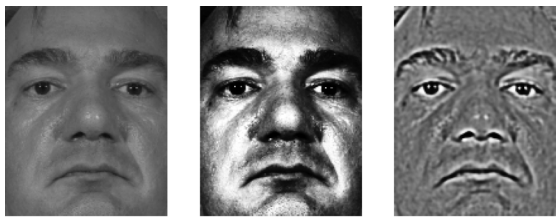


Fig. 14. Preprocessed face images with different photometric normalization methods. (a) "NO". (b) "HE". (c) "PP".

Fig. 14 illustrates one example of the preprocessed images using these methods.

From Table IV, we can find that neither "HE" nor "PP" can consistently improve the performance of all these methods for both Experiment 1 and 4. Specifically, by using the "HE" normalization, all the methods achieve the best results in Experiment 1, but in Experiment 4 they perform even worse than those with "NO" photometric normalization. In contrast, by using the "PP" method, all the methods except "LGXP+BFLD" perform the best in Experiment 4 but just comparable to the results with "NO" photometric normalization. Another important observation is that all the methods with "NO" illumination preprocessing achieve results comparable to the best ones with the "HE" or "PP" normalization.

From these comparisons, we might draw two conclusions: 1) by carefully choosing appropriate photometric normalization methods, the methods can be further improved; 2) the proposed method by itself is very robust to illumination variations. We attribute the robustness to two points. First, the local patterns of Gabor feature are robust to the illumination variations. Specifically, it is known that Gabor magnitude feature is robust to the illumination variations, and the proposed quantized Gabor phase codes also result in the robustness by using the large-step quantizing operation. Second, the BFLD method leads to robustness to illumination variations, since the BFLD can model these variations as intrapersonal variations.

Besides the above comparison, we also compare the results of our method with those results reported recently on this database with the same evaluation protocol, as shown in Table V. In the table, all the results for comparison are directly cited from the related papers. For clarity, only the results of ROC3 are listed in the table. From the comparison, we can see that our score-level fusion method works comparatively well to the best known results reported in [30], which further validates the effectiveness of the proposed techniques.

VII. CONCLUSION

This paper investigates how to exploit effectively the Gabor phase information, as well as its fusion with Gabor magnitude. Under the framework of local pattern encoding, we propose the so-called Local Gabor XOR Patterns (LGXP) to encode locally the Gabor phase. Additionally, by one " *divide and conquer*" strategy, we introduce the BFLD method, and study

TABLE IV
VERIFICATION RATES ON FRGC 2.0 DATABASE WITH DIFFERENT PHOTOMETRIC NORMALIZATION METHODS WHEN $\mathrm{FAR} = 0.1\%$

Methods	Photometric normalization	Exp.1			Exp.4		
		ROC1	ROC2	ROC3	ROC1	ROC2	ROC3
LGBP_Mag+BFLD	NO	98.1%	97.5%	96.8%	79.6%	80.3%	81.0%
	HE	**98.2%**	**97.6%**	**97.0%**	73.4%	74.3%	74.8%
	PP	98.0%	97.4%	96.6%	**80.1%**	**80.7%**	**81.3%**
LGXP+BFLD	NO	98.5%	97.9%	97.3%	**78.2%**	**78.6%**	**78.9%**
	HE	**98.7%**	**98.0%**	**97.3%**	70.3%	70.5%	70.7%
	PP	98.4%	97.6%	96.9%	78.0%	78.4%	78.6%
F[LGBP_Mag+LGXP]	NO	98.6%	98.0%	97.3%	81.9%	82.6%	83.1%
	HE	**98.7%**	**98.1%**	**97.5%**	75.9%	76.6%	77.1%
	PP	98.6%	97.9%	97.2%	**82.7%**	**83.2%**	**83.6%**
S[LGBP_Mag+LGXP]	NO	98.7%	98.1%	97.5%	83.6%	84.3%	84.9%
	HE	**98.8%**	**98.3%**	**97.8%**	78.7%	79.4%	80.1%
	PP	98.6%	98.0%	97.3%	**83.9%**	**84.7%**	**85.2%**

TABLE V
COMPARISONS WITH SOME STATE-OF-THE-ART METHODS ON FRGC 2.0 DATABASE (VR WHEN $\mathrm{FAR} = 0.1\%$)

Methods	Exp.1	Exp.4
	ROC3	ROC3
Hwang et al's Method in [12]	91.5%	74.3%
Liu's Method in [20]	92.0%	76.0%
Liu and Liu's Method in [21]	N/A	81.3%
Su et al's Method in [30]	**98.0%**	**86.0%**
Tan and Triggs's Method in [32]	N/A	83.6%
Proposed score-level fusion method	97.5%	84.9%

Note: As is in this work, the eye positions in original face images are also manually located for all the comparison methods in [12], [20], [21], [30] and [32].

the fusion methods of different local patterns of Gabor feature. These methods are extensively evaluated and compared with previous methods on the FERET and FRGC 2.0 datasets, which indicates that our fusion method achieves better or comparable results than the best known ones.

Unlike previous work, this study indicates that Gabor phase might embody more (at least equal) discriminating power than Gabor magnitude, if only it is appropriately exploited. We observe this phenomenon on both FERET datasets and Experiments 1 and 4 of FRGC 2.0.

We also experimentally reveal that methods based on local Gabor patterns, both the proposed LGXP and existing LGBP/LGPP, work reasonably well under relatively simple testing scenarios, for instance, FERET Fb, Fc, and Experiment 1 of FRGC 2.0. But, they all degrade abruptly when the testing is challenging with large variations due to unconstrained imaging conditions, e.g., FERET DupI, DupII, especially Experiment 4 of FRGC 2.0. This observation might imply that these local pattern methods are sensitive to large extrinsic variations.

Fortunately, the challenging problems caused by unconstrained conditions can be well addressed by combining local Gabor patterns with the proposed BFLD method. Impressive improvements are achieved on FERET DupI, DupII, and Experiment 4 of FRGC 2.0, which safely validate the effectiveness of BFLD for challenging testing scenarios. We attribute this performance gain to the discriminating nature of FLD and the "divide and conquer" methodology brought by the block-based strategy.

Finally, as expected, the fusion of magnitude and phase further enhance the recognition accuracy when they are encoded by local patterns and combined with BFLD. In the current score-level fusion, we simply combine component classifiers by sum rule. It is evidently not optimal, and as a future work we will exploit better statistical fusion schemes. Additionally, as one statistical learning method, the generalization ability of BFLD method is influenced greatly by the training set. We need more future efforts to improve its performance for out-of-sample problem.

ACKNOWLEDGMENT

The authors would like to thank Prof. A. Ross and the anonymous reviewers, whose comments helped to improve the paper greatly. They would also like to thank Dr. B. Zhang and Dr. W. Zhang for their help on this work. Portions of the research in this paper use the FERET database of facial images collected under the FERET program, sponsored by the DOD Counterdrug Technology Development Program Office. They would also thank P. J. Philips and other FRGC members for their contributions(http://www.frvt.org/FRGC/).

REFERENCES

[1] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.

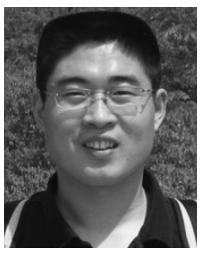
[2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[3] R. Brunelli and T. Poggio, "Face recognition: Features versus templates," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 10, pp. 1042–1052, Oct. 1993.

[4] D. Cai, X. He, J. Han, and H. Zhang, "Orthogonal Laplacianfaces for face recognition," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3608–3614, Nov. 2006.

[5] J. T. Chien and C. C. Wu, "Discriminant waveletfaces and nearest feature classifiers for face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 12, pp. 1644–1649, Dec. 2002.

[6] J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Amer. A*, vol. 2, no. 7, pp. 1160–1169, 1985.
[7] J. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 11, pp. 1148–1161, Nov. 1993.
[8] Y. Gao, Y. Wang, X. Zhu, X. Feng, and X. Zhou, "Weighted Gabor features in unitary space for face recognition," presented at the Int. Conf. Automatic Face and Gesture Recognition, 2006.
[9] 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.
[10] X. He, S. Yan, Y. Hu, P. Niyogi, and H. Zhang, "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.
[11] B. Heisele, P. Ho, and T. Poggio, "Face recognition with support vector machines: Global versus component-based approach," in *Proc. Int. Conf. Computer Vision*, 2001, pp. 688–694.
[12] W. Hwang, G. Park, and J. Lee, "Multiple face model of hybrid Fourier feature for large face image set," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2006, pp. 1574–1581.
[13] J. P. Jones and L. A. Palmer, "An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex," *J. Neurophysiol.*, vol. 58, no. 6, pp. 1233–1258, 1987.
[14] T. Kim, H. Kim, W. Hwang, and J. Kittler, "Component-based LDA face description for image retrieval and MPEG-7 standardisation," *Image Vis. Comput.*, pp. 631–642, 2005.
[15] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 226–239, Mar. 1998.
[16] M. Lades, J. C. Vorbrüggen, J. Buhmann, J. Lange, C. v. d. Malsburg, R. P. Würtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. Computers*, vol. 42, no. 3, pp. 300–311, 1993.
[17] J. Lai, P. C. Yuen, and G. Feng, "Face recognition using holistic Fourier invariant features," *Pattern Recognit.*, vol. 34, no. 1, pp. 95–109, 2001.
[18] D. Lin and X. Tang, "Recognize high resolution faces: From macrocosm to microcosm," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2006, pp. 1355–1362.
[19] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 467–476, Apr. 2002.
[20] C. Liu, "Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 725–737, May 2006.
[21] Z. Liu and C. Liu, "Fusion of the complementary Discrete Cosine Features in the YIQ color space for face recognition," *Comput. Vis. Image Understand.*, vol. 111, no. 3, pp. 249–262, 2008.
[22] B. Moghaddam, T. Jebara, and A. Pentland, "Bayesian face recognition," *Pattern Recognit.*, vol. 33, pp. 1771–1782, 2000.
[23] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
[24] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 10, pp. 1090–1104, Oct. 2000.
[25] P. J. Phillips, P. J. Flynn, T. Scruggs, K. W. Bowyer, J. Chang, K. Hoffman, J. Marques, J. Min, and W. Worek, "Overview of the face recognition grand challenge," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, 2005, pp. 947–954.
[26] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 12, pp. 2323–2326, 2000.
[27] M. Savvides, B. V. K. Vijaya Kumar, and P. K. Khosla, "Eigenphases vs. Eigenfaces," in *Proc. Int. Conf. Pattern Recognition*, 2004, pp. 810–813.
[28] S. Shan, W. Zhang, Y. Su, X. Chen, and W. Gao, "Ensemble of piecewise FDA based on spatial histograms of local (Gabor) binary patterns for face recognition," in *Proc. Int. Conf. Pattern Recognition*, 2006, pp. 590–593.
[29] L. Shen and L. Bai, "A review on Gabor wavelets for face recognition," *Pattern Anal. Appl.*, vol. 9, no. 10, pp. 273–292, 2006.
[30] Y. Su, S. Shan, X. Chen, and W. Gao, "Hierarchical ensemble of global and local classifiers for face recognition," presented at the Int. Conf. Computer Vision, 2007.
[31] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," presented at the IEEE Int. Workshop on Analysis and Modeling of Faces and Gestures, 2007.
[32] X. Tan and B. Triggs, "Fusing Gabor and LBP feature sets for kernel-based face recognition," presented at the IEEE Int. Workshop on Analysis and Modeling of Faces and Gestures, 2007.
[33] D. Tao, X. Li, X. Wu, and S. J. Maybank, "Geometric mean for subspace selection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 260–274, Feb. 2009.
[34] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 12, pp. 2319–2322, 2000.
[35] M. Turk and A. Pentland, "Eigenfaces for recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
[36] X. Wang and X. Tang, "Dual-space linear discriminant analysis for face recognition," presented at the IEEE Computer Society Conf. Computer Vision and Pattern Recognition, 2004.
[37] L. Wiskott, J.-M. Fellous, N. Krüger, and C. von der Malsburg, "Face recognition by Elastic Bunch graph matching," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 775–779, Jul. 1997.
[38] X. Xie and K.-M. Lam, "Gabor-based kernel PCA with doubly nonlinear mapping for face recognition with a single face image," *IEEE Trans. Image Process.*, vol. 15, no. 9, pp. 2481–2492, Sep. 2006.
[39] S. Yan, D. Xu, B. Zhang, H. Zhang, Q. Yang, and S. Lin, "Graph embedding and extensions: A general framework for dimensionality reduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 1, pp. 40–51, Jan. 2007.
[40] J. Yang, D. Zhang, A. F. Frangi, and J. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131–137, Jan. 2004.
[41] J. Yang, D. Zhang, J. Yang, and B. Niu, "Globally maximizing, locally minimizing: Unsupervised discriminant projection with applications to face and palm biometrics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 650–664, Apr. 2007.
[42] P. Yang, S. Shan, W. Gao, S. Z. Li, and D. Zhang, "Face recognition using Ada-boosted Gabor features," in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition*, 2004, pp. 356–361.
[43] B. Zhang, S. Shan, X. Chen, and W. Gao, "Histogram of Gabor Phase Patterns (HGPP): A novel object representation approach for face recognition," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 57–68, Jan. 2007.
[44] D. Zhang, W.-K. Kong, J. You, and M. Wong, "Online palmprint identification," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 9, pp. 1041–1050, Sep. 2003.
[45] L. Zhang, S. Z. Li, Z. Qu, and X. Huang, "Boosting local feature based classifiers for face recognition," presented at the IEEE Computer Society Conf. Computer Vision and Pattern Recognition Workshops, 2004.
[46] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local Gabor Binary Pattern Histogram Sequence (LGBPHS): A novel non-statistical model for face representation and recognition," in *Proc. Int. Conf. Computer Vision*, 2005, pp. 786–791.
[47] W. Zhang, S. Shan, X. Chen, and W. Gao, "Are Gabor phases really useless for face recognition?," in *Proc. Int. Conf. Pattern Recognition*, 2006, pp. 606–609.
[48] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: A literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, 2003.
[49] J. Zou, Q. Ji, and G. Nagy, "A comparative study of local matching approach for face recognition," *IEEE Trans. Image Process.*, vol. 16, no. 10, pp. 2617–2628, Oct. 2007.



Shufu Xie received the B.S. and M.S. degrees in computer science from the School of Information Science and Engineering, Shandong Normal University, Jinan, China, in 2003 and 2006, respectively. Currently, he is pursuing the Ph.D. degree at the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China.

His research interests include pattern recognition, image processing, and especially focus on face recognition.

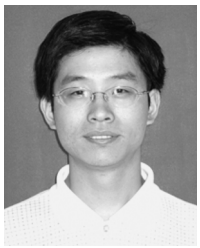
Shiguang Shan (M'04) received the M.S. degree in computer science from the Harbin Institute of Technology, Harbin, China, in 1999, and the Ph.D. degree in computer science from the Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS), Beijing, in 2004.

He has been with ICT, CAS, since 2002 and has been an Associate Professor since 2005. He is also the Vice Director of the ICT-ISVISION Joint Research and Development Laboratory for Face Recognition, ICT, CAS. His research interests cover image analysis, pattern recognition, and computer vision. He is focusing especially on face recognition related research topics, and has published more than 100 papers on the related research topics. He received the China's State Scientific and Technological Progress Awards in 2005 for his work on face recognition technologies.

Xilin Chen (M'00–SM'09) received the B.S., M.S., and Ph.D. degrees in computer science from the Harbin Institute of Technology, Harbin, China, in 1988, 1991, and 1994, respectively.

He was a Professor with the Harbin Institute of Technology from 1999 to 2005. He was a Visiting Scholar with Carnegie Mellon University, Pittsburgh, PA, from 2001 to 2004. He has been with the Institute of Computing Technology, Chinese Academy of Sciences (CAS), Beijing, since August 2004. He is the Director of the Key Laboratory of Intelligent Information Processing, CAS. He has published one book and over 150 papers in refereed journals and proceedings in the areas of computer vision, pattern recognition, image processing, and multimodal interfaces.

Dr. Chen has served as a program committee member for more than 30 international and national conferences. He has received several awards, including China's State Scientific and Technological Progress Award in 2000, 2003, and 2005 for his research work.



Jie Chen (M'09) received the M.S. and Ph.D. degrees from the Harbin Institute of Technology, Harbin, China, in 2002 and 2007, respectively.

Since September 2007, he has been a senior researcher in the Machine Vision Group at the University of Oulu, Finland. His research interests include pattern recognition, computer vision, machine learning, dynamic texture, and watermarking. He has authored over 20 papers in journals and conferences.