Histogram of Gabor Phase Patterns (HGPP): A Novel Object Representation Approach for Face Recognition

Baochang Zhang, Shiguang Shan, *Member, IEEE*, Xilin Chen, *Member, IEEE*, and Wen Gao, *Senior Member, IEEE*

***Abstract*—A novel object descriptor, histogram of Gabor phase pattern (HGPP), is proposed for robust face recognition. In HGPP, the quadrant-bit codes are first extracted from faces based on the Gabor transformation. Global Gabor phase pattern (GGPP) and local Gabor phase pattern (LGPP) are then proposed to encode the phase variations. GGPP captures the variations derived from the orientation changing of Gabor wavelet at a given scale (frequency), while LGPP encodes the local neighborhood variations by using a novel local XOR pattern (LXP) operator. They are both divided into the nonoverlapping rectangular regions, from which spatial histograms are extracted and concatenated into an extended histogram feature to represent the original image. Finally, the recognition is performed by using the nearest-neighbor classifier with histogram intersection as the similarity measurement. The features of HGPP lie in two aspects: 1) HGPP can describe the general face images robustly without the training procedure; 2) HGPP encodes the Gabor phase information, while most previous face recognition methods exploit the Gabor magnitude information. In addition, Fisher separation criterion is further used to improve the performance of HGPP by weighing the subregions of the image according to their discriminative powers. The proposed methods are successfully applied to face recognition, and the experiment results on the large-scale FERET and CAS-PEAL databases show that the proposed algorithms significantly outperform other well-known systems in terms of recognition rate.**

***Index Terms*—Face recognition, feature extraction, Gabor, histogram, local pattern, phase pattern.**

I. Introduction

THE intensity of an image is the only source from a camera used for object recognition. However, a lot of variations, such as albedo and shape of the object, lighting, etc., are all encoded as the intensity. To eliminate the extrinsic factors, various feature extraction and selection methods are widely used [1]–[4].

Manuscript received November 16, 2005, revised July 7, 2006. The work was sponsored in part by "100 Talents Program" of CAS, in part by the Natural Science Foundation of China under Contract 60332010, and in part by ISVISION Technologies, Co., Ltd. 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. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Ljubisa Stankovic.

B. Zhang is with the Harbin Institute of Technology, Harbin 150001, China (e-mail: bczhang@ jdl.ac.cn).

S. Shan and X. Chen are with Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China (e-mail: sgshan@jdl.ac.cn; xlchen@jdl.ac.cn).

W. Gao is with the Harbin Institute of Technology, Harbin 150001, China, and also with Peking University, Beijing 100871, China, and the Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, China (e-mail: wgao@jdl.ac.cn).

Digital Object Identifier 10.1109/TIP.2006.884956

To achieve the goal of extracting features in a certain scale, Laplacian of Gaussian (LoG) was first introduced by Marr and Hildreth [5] to simulate the lateral inhibition for edge detection. However, to be a powerful descriptor, the feature extractor should be anisotropic, which means that it should enhance the feature in a certain scale and orientation simultaneously. In recent years, the Gabor transformation has been widely used as an effective element in the image processing and pattern recognition tasks. The Gabor wavelet [6] was first introduced by David Gabor in 1946. The Gabor wavelet is a sinusoidal plane wave with a particular frequency and orientation, modulated by a Gaussian envelope. It can characterize the spatial frequency structure in the image while preserving information of spatial relations and, thus, is suitable for extracting the orientation-dependent frequency contents of patterns.

Lades *et al.* [7] pioneered the use of the Gabor wavelet for face recognition using the dynamic link architecture framework. Wiskott *et al.* [8] subsequently developed a Gabor wavelet-based elastic bunch graph matching (EBGM) method to label and recognize human faces. In the EBGM method, the face is represented as a graph, each node of which contains a group of coefficients, known as a *jet*. It can also measure the geometry of the face by using the labeled distance vector, which is the edge part of the graph. Liu and Wechsler [9] showed that the face representation based on the magnitude part of Gabor feature had been a promising way towards achieving high accuracy face recognition. Shan *et al.* [10] provided an AdaBoost-based strategy to select the discriminative features from the magnitude part of the Gabor feature, and then trained a Fisher classifier to make a final classification. As a powerful descriptor, the Gabor wavelet is also used in many applications, such as data compression [11], optical character recognition (OCR) [12], texture analysis [13], fingerprint recognition [14], and so on. Most of the above applications are based on the magnitude part of Gabor feature. In fact, the Gabor phase is a very discriminative information source, and has been successfully used in iris and palmprint identification [15], [16].

The histogram method has been widely used to represent, analyze, and recognize images because it can be calculated easily and efficiently, and robust to the noise and local image transformations [17]–[19]. One of initial applications of histograms was the work of Swain and Ballard for the identification of 3-D objects [17]. Subsequently, various recognition systems [18]–[20] based on histograms were developed. However, the histogram is not adequate for many applications as it suffers from losing the structure information of the object. Multiresolution histograms have been proposed to encode the structure information by convolving the image with Gaussian filters

[20], difference of Gaussians [21], Gabor filters [22], and so on. Spatial histograms [23], [24] aim to capture the spatial relationships between parts of images. Thus, it is more suitable for object recognition, such as face recognition. Ahonen *et al.* [23] presented a novel approach based on Local Binary Pattern (LBP) histograms for face recognition, which considered both shape and texture information in representing face images. The idea behind using the LBP features is that the face image can be seen as a composition of micro-patterns [23]. Zhang *et al.* [22] proposed the so-called local Gabor binary pattern histogram sequence (LGBPHS) method, which combined the magnitude part of Gabor feature and the LBP operator. The LGBPHS method proposed by our group achieved a better result than other well-known face recognition systems on the standard FERET database. Both methods are, in nature, based on the spatial histogram, which can capture the structure information of the input face object and provide an easy matching strategy.

In this paper, we propose a compact and effective object descriptor – histogram of Gabor phase pattern (HGPP) for robust face recognition. The approach is based on the combination of the spatial histogram and the Gabor phase information encoding scheme, the so-called Gabor phase pattern (GPP). Different from the learning-based face recognition methods, in HGPP, features are directly extracted without the training procedure. In this paper, two kinds of GPPs, GGPP and LGPP, are proposed to capture the phase variations derived from the orientation changing of Gabor wavelet and the relationships among local neighbors. They are both divided into small nonoverlapping rectangular regions, from which the local histograms are extracted and concatenated into a single extended histogram feature to capture the spatial information. The intuition behind HGPP is that the proposed GGPP or LGPP is also a kind of method modeling micro-patterns. Motivated by LBP-based micro-pattern [22], [23] methods, we used the spatial histogram to preserve the information about the distribution of local micro-patterns, such as edges, spots, and flat areas [23]. Another reason for the combination of Gabor wavelets and histogram lies in the observation that GPP contains rich texture information and histogram serves as a good description tool for texture images. The recognition is performed using the nearest-neighbor classifier with histogram intersection as the similarity measurement. We also make further improvement by using the Fisher separation criterion (FSC) to weight the subregions of the object according to their discriminative powers.

We have applied our method to the face recognition problem [25], [26], which is an active topic in the computer vision research. Face recognition can be widely used in real environments, such as surveillance and security, telecommunication, digital libraries, human computer interaction (HCI), and smart environments. Experiments on the large-scale FERET [27] and CAS-PEAL [28] databases have been performed to evaluate the effectiveness of the proposed method, which include testing its performance against different facial expression, lighting, and aging of subjects. Experimental results show that the proposed method impressively outperforms other well-known systems.

The rest of the paper is organized as follows. In Section II, HGPP is proposed to make feature extraction from the original image. In Section III, we propose the basic HGPP and the weighted HGPP matching methods based on histogram intersection. In Section IV, the proposed method is applied to the face recognition problem. Two groups of experiments on the large-scale FERET and CAS-PEAL databases are conducted to evaluate its performance. In the last section, some brief conclusions are drawn with some discussion on the future work.

II. HGPP: A Novel Object Representation Approach

In this section, we propose two schemes, GGPP and LGPP, to encode the Gabor phase information. Both of them are defined based on the quadrant-bit codes of Gabor real and imaginary parts, $(P^{\mathrm{Re}}_{u,v}(Z), P^{\mathrm{Im}}_{u,v}(Z))$, proposed by Daugman for iris recognition [17]. While the proposed GGPP further encodes the orientation information at each scale, LGPP encodes the local neighborhood variations at each orientation and scale. Finally, GGPP and LGPP are combined with spatial histograms to model the original object image.

A. Quadrant Bit Coding (QBC)

The Gabor wavelets (kernels, filters) can be defined as follows [9]:

$$\psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{(-\|k_{u,v}\|^2\|z\|^2/2\sigma^2)} \left[e^{ik_{u,v}z} - e^{-\sigma^2/2} \right] \tag{1}$$

where $\overrightarrow{k_{u,v}} = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_v \cos\phi_u \\ k_v \sin\phi_u \end{pmatrix}$, $k_v = f_{\max}/2^{v/2}$, $\phi_u = u(\pi/8)$, $v = 0, \ldots, v_{\max} - 1$, $u = 0, \ldots, u_{\max} - 1$, v is the frequency and u is the orientation with $v_{\max} = 5$, $u_{\max} = 8$, $\sigma = 2\pi$. Each kernel is a product of a Gaussian envelope and a complex plane wave, while the first term in the square brackets in (1) determines the oscillatory part of the kernel and the second term compensates for the DC value. In addition, σ determines the ratio of the Gaussian window width to wavelength [9]. The Gabor transformation of a given image is defined as its convolution with the Gabor functions

$$G_{u,v}(z) = I(z) * \Psi_{u,v}(z) \tag{2}$$

where $z = (x, y)$ denotes the image position, the symbol "$*$" denotes the convolution operator, and $G_{u,v}(z)$ is the convolution result corresponding to the Gabor kernel at scale v and orientation u. The Gabor wavelet coefficient $G_{u,v}(z)$ is a complex, which can be rewritten as [8]

$$G_{u,v}(z) = A_{u,v}(z) \cdot \exp(i\theta_{u,v}(z)) \tag{3}$$

with one magnitude item $A_{u,v}(z)$, and one phase item $\theta_{u,v}(z)$. It is known that the magnitude varies slowly with the spatial position, while the phases rotate in some rate with positions, as can be seen from the examples in Fig. 1. Due to this rotation, the phases taken from image points only a few pixels apart have very different values, although representing almost the same local feature [8]. This can cause severe problems for object (face) matching, and it is just the reason that most previous works make use of only the magnitude for face classification.

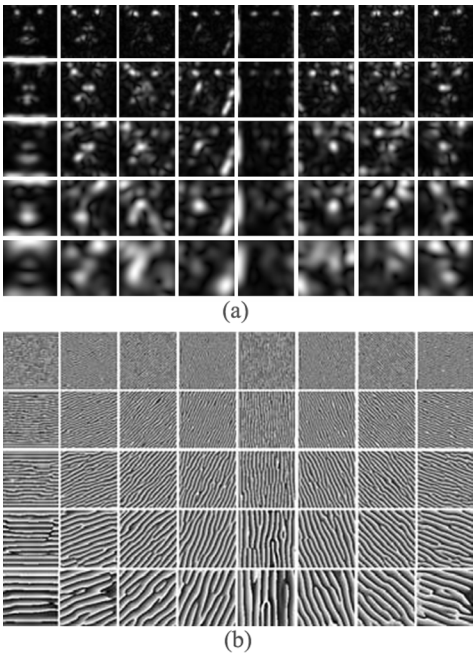

Fig. 1. Visualization of the Gabor phase and magnitude. (a) Gabor magnitude. (b) Gabor phase.

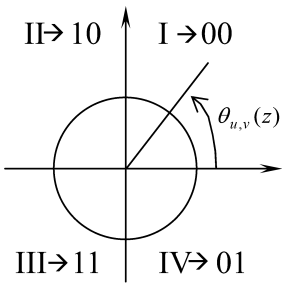

Fig. 2. Quadrant bit coding of Gabor phase.

However, the Gabor phase is not worthless. A typical successful application of Gabor phase is the phase-quadrant demodulation coding method proposed by Daugman for iris recognition [15]. In [15], each pixel in the resultant image will be encoded to two bits, $(P^{\mathrm{Re}}_{u,v}(Z), P^{\mathrm{Im}}_{u,v}(Z))$, by the following rules:

$$P^{\mathrm{Re}}_{u,v}(Z) = \left\{ \begin{array}{ll} 0, & \text{if } \mathrm{Re}(G_{u,v}(Z)) > 0 \\ 1 & \text{if } \mathrm{Re}(G_{u,v}(Z)) <= 0 \end{array} \right\} \quad (4)$$

$$P^{\mathrm{Im}}_{u,v}(Z) = \left\{ \begin{array}{ll} 0, & \text{if } \mathrm{Im}(G_{u,v}(Z)) > 0 \\ 1, & \text{if } \mathrm{Im}(G_{u,v}(Z)) <= 0 \end{array} \right\} \quad (5)$$

where $\mathrm{Re}(G_{u,v}(Z))$ and $\mathrm{Im}(G_{u,v}(Z))$ are respectively the real and imaginary parts of the Gabor coefficient.

Daugman's encoding method, given by (4) and (5), can be reformulated as follows:

$$P^{\mathrm{Re}}_{u,v}(Z) = \left\{ \begin{array}{ll} 0, & \text{if } \theta_{u,v}(Z) \in \{I, IV\} \\ 1, & \text{if } \theta_{u,v}(Z) \in \{II, III\} \end{array} \right\} \quad (6)$$

$$P^{\mathrm{Im}}_{u,v}(Z) = \left\{ \begin{array}{ll} 0, & \text{if } \theta_{u,v}(Z) \in \{I, II\} \\ 1, & \text{if } \theta_{u,v}(Z) \in \{III, IV\} \end{array} \right\} \quad (7)$$

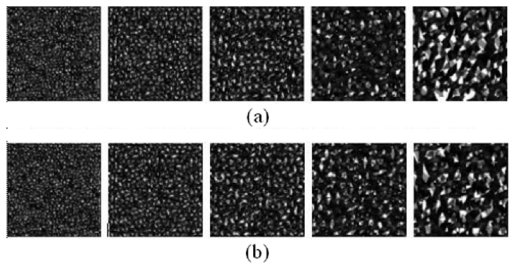

Fig. 3. (a) Visualization of real part GGPP (b) Visualization of imaginary part GGPP.

where $\theta_{u,v}(Z)$ is the Gabor phase angle for the pixel at the position Z. It is obvious that (6) and (7) [(4) and (5)], the so-called quadrant bit coding (QBC), assign two bits for each pixel according to the quadrant in which the Gabor phase angle lies. QBC is relatively stable, and it is actually the quantification of Gabor feature. It results in the same feature ("00") for the phase angles in $(0°, 90°]$ (similarly for $(90°, -180°]$, $(180°, -°270°]$, $(270°, -°360°]$). This should be one of reasons why it is successfully applied to the Iris and Palmprint recognition problems. This encoding procedure can be understood more clearly as shown in Fig. 2.

B. Global Gabor Phase Patterns

Simply speaking, the GGPP scheme computes one binary string for each pixel by concatenating the real (or imaginary) quadrant-bit codes of different orientations for a given frequency. Formally, the GGPP value, $GGPP_v(Z_0)$, for the frequency v at the position Z_0 in a given image is formulated as

$$GGPP^{\mathrm{Re}}_v(Z_0) = [P^{\mathrm{Re}}_{0,v}(Z_0), P^{\mathrm{Re}}_{1,v}(Z_0), \ldots, P^{\mathrm{Re}}_{k,v}(Z_0)] \quad (8)$$

$$GGPP^{\mathrm{Im}}_v(Z_0) = [P^{\mathrm{Im}}_{0,v}(Z_0), P^{\mathrm{Im}}_{1,v}(Z_0), \ldots, P^{\mathrm{Im}}_{k,v}(Z_0)] \quad (9)$$

where (8) gives the real GGPP and (9) is the imaginary GGPP. In our paper, eight orientations are exploited, i.e., $\mathrm{k} = 7$, which "happens" to form a byte to represent 256 different orientation modes. These modes can be easily computed by the following equations:

$$\begin{aligned} GGPP^{\mathrm{Re}}_v(Z_0) = & P^{\mathrm{Re}}_{0,v}(Z_0) * 2^k \\ & + P^{\mathrm{Re}}_{1,v}(Z_0) * 2^{k-1} + \cdots + P^{\mathrm{Re}}_{k,v}(Z_0)) \quad (10) \\ GGPP^{\mathrm{Im}}_v(Z_0) = & P^{\mathrm{Im}}_{0,v}(Z_0) * 2^k \\ & + P^{\mathrm{Im}}_{1,v}(Z_0) * 2^{k-1} + \cdots + P^{\mathrm{Im}}_{k,v}(Z_0)) \quad (11) \end{aligned}$$

where (10) and (11) are used to calculate the decimal values for the 8-bit binary strings, i.e., (8) and (9). By using this encoding method, one can get two decimal numbers for each pixel corresponding to the real and imaginary GGPPs. Both of them range in [0, 255], and it is easy to visualize them as the grey-level images for a given frequency. Examples of the real and imaginary GGPPs are shown in Fig. 3(a) and (b) for intuitive understandings. They are also micro-patterns, and look like maps with rich texture, which motivate us to combine them with the histogram features.

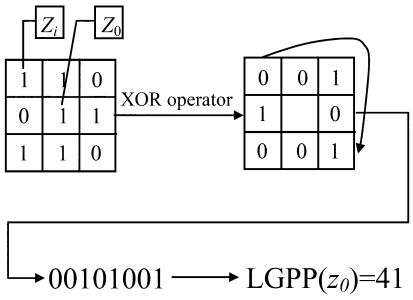

Fig. 4. Illustration of the LXP operator.

C. Local Gabor Phase Patterns

Unlike GGPP encoding orientation information, we further proposed to encode the local variations for each pixel, which we call LGPP. Formally, for each orientation u and frequency v, the real or imaginary LGPP value at the position Z_0 is computed through the following equation, named local XOR pattern (LXP) operator:

$$\begin{aligned} &LGPP^{\mathrm{Re}}_{u,v}(Z_0) \\ &= [P^{\mathrm{Re}}_{u,v}(Z_0)_{\mathrm{XOR}}P^{\mathrm{Re}}_{u,v}(Z_1) \\ &\quad P^{\mathrm{Re}}_{u,v}(Z_0)_{\mathrm{XOR}}P^{\mathrm{Re}}_{u,v}(Z_2),\ldots,P^{\mathrm{Re}}_{u,v}(Z_0)\mathrm{XOR}P^{\mathrm{Re}}_{u,v}(Z_8)] \end{aligned} \tag{12}$$

$$\begin{aligned} &LGPP^{\mathrm{Im}}_{u,v}(Z_0) \\ &= [P^{\mathrm{Im}}_{u,v}(Z_0)_{\mathrm{XOR}}P^{\mathrm{Im}}_{u,v}(Z_1) \\ &\quad P^{\mathrm{Im}}_{u,v}(Z_0)_{\mathrm{XOR}}P^{\mathrm{Im}}_{u,v}(Z_2),\ldots,P^{\mathrm{Im}}_{u,v}(Z_0)_{\mathrm{XOR}}P^{\mathrm{Im}}_{u,v}(Z_8)] \end{aligned} \tag{13}$$

where Z_i, $i = 1, 2, \ldots, 8$, are the eight neighbors around Z_0, and XOR denotes the bit *exclusive or* operator. This procedure is illustrated in Fig. 4 for easy understanding.

By recalling the definition of QBC [(4) and (5)], the computation of each bit in (12) and (13) is actually equivalent to

$$\begin{aligned} &P^{\mathrm{Re}}_{u,v}(Z_0)_{\mathrm{XOR}}P^{\mathrm{Re}}_{u,v}(Z_i) \\ &= \begin{cases} 0, & \text{if } \mathrm{Re}(\mathrm{G}_{u,v}(Z_0)) * \mathrm{Re}(\mathrm{G}_{u,v}(Z_i)) > 0 \\ 1, & \text{if } \mathrm{Re}(\mathrm{G}_{u,v}(Z_0)) * \mathrm{Re}(\mathrm{G}_{u,v}(Z_i)) <= 0 \end{cases} \end{aligned} \tag{14}$$

$$\begin{aligned} &P^{\mathrm{Im}}_{u,v}(Z_0)_{\mathrm{XOR}}P^{\mathrm{Im}}_{u,v}(Z_1) \\ &= \begin{cases} 0, & \text{if } \mathrm{Im}(\mathrm{G}_{u,v}(Z_0)) * \mathrm{Im}(G_{u,v}(Z_i)) > 0 \\ 1, & \text{if } \mathrm{Im}(\mathrm{G}_{u,v}(Z_0)) * \mathrm{Im}(G_{u,v}(Z_i)) <= 0 \end{cases}. \end{aligned} \tag{15}$$

From (12) and (13), one can clearly see that LGPP actually encodes the *sign* difference of the central pixel from its neighbors. Therefore, LGPP can also reveal the spots ("111111111"), flat areas ("00000000"), and so on, for binary images. One can also refer to LBP [29].

Similar to GGPP, eight neighbors again "happen" to provide 8 bits to form a byte for each pixel. Therefore, a decimal number ranged in [0, 255] can be computed as (10) and (11). Each value represents a mode how the pixel at the position Z_0 is different from its 8 neighbors. This procedure will generate $v_{\max} * u_{\max} * 2$ LGPP "images" for the real and imaginary LGPPs of all orientations and scales (frequencies). Examples of the images are illustrated in Fig. 5, from which one can also find the rich structural *fingerprint-like* texture. Other samples of LGPPs are given in Fig. 6, which shows that different people have different Gabor phase patterns.

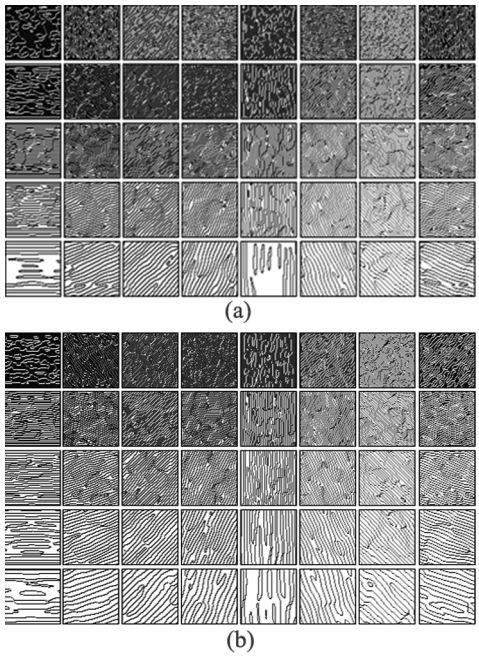

Fig. 5. (a) Visualization of real part LGPP. (b) Visualization of imaginary part LGPP.

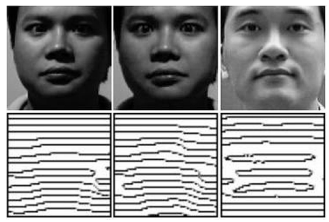

Fig. 6. $\mathrm{LGPP}_{0,4}(\mathrm{Z})$ for three images.

In [22], the LGBPHS method is proposed for face recognition, which is also based on Gabor feature and the spatial histogram. The main differences between LGBPHS and LGPP are as follows.

1) The LGBPHS method is based on the magnitude part of Gabor feature, while LGPP exploits the Gabor phase information.
2) The LGBPHS method is an application of the LBP operator, which is a powerful means of the texture descriptor introduced by Ojala [29]. While in LGPP, a novel LXP operator is proposed to capture the relationships among local neighbors as to whether they lie in the same quadrants. Just like the LBP features, the LGPP features can also be a kind of method modeling micro-patterns, by which a description of the face image is obtained.

D. Histogram of Gabor Phase Patterns

In Daugman's iris recognition method, quadrant-bit codes are directly used to form the representation of an iris image

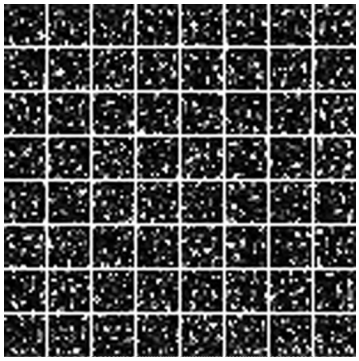

Fig. 7. $\mathrm{LGPP}_0(\mathrm{Z})$ divided into 64 subregions.

and the classification is achieved by hamming distance. However, for five scales and eight orientations, the proposed GPP encoding method will get 90 "images" (five real GGPPs, five imaginary GGPPs, 40 real LGPPs, and 40 imaginary LGPPs), with the same size as the original face image. To model them more efficiently and compactly, in this paper, we further exploit the spatial histogram to model the encoded GPPs (GGPP and LGPP). The reason that we turn to histogram-based approaches lies in the observations that the GPPs, both GGPP and LGPP micro-patterns, look like the images with rich structural textures (refer to Figs. 3 and 5), and histogram serves as a good description tool for representing the texture images [20], [22].

However, a single global histogram suffers from losing the structure information of the object. In the case of face recognition, a face is a topological object, which means that the spatial structure is of high importance for the face recognition task. In order to reserve the spatial information in the histogram features, the GPP images are spatially divided into the nonoverlapping rectangular regions (as the example shown in Fig. 7), represented by $R_1, \ldots, R_L$, from which the spatial histograms are extracted respectively. Then, all of these histograms are concatenated into a single extended histogram feature, the so-called joint local-histogram feature (JLHF), for all frequencies and orientations. We call the resulting representation, i.e., JLHF of the GPP images, HGPP.

Formally, the HGPP feature is formulated as

$$HGPP = (H^{\mathrm{Re}}_{GGPP}, H^{\mathrm{Im}}_{GGPP}, H^{\mathrm{Re}}_{LGPP}, H^{\mathrm{Im}}_{LGPP}) \qquad (16)$$

where $H^{\mathrm{Re}}_{GGPP}, H^{\mathrm{Im}}_{GGPP}, H^{\mathrm{Re}}_{LGPP}, H^{\mathrm{Im}}_{LGPP}$ are the subregion histograms of the real part GGPP, the subregion histograms of the imaginary part GGPP, the subregion histograms of the real part LGPP, and the subregion histograms of the imaginary part LGPP, respectively, and formulated as

$$\begin{aligned}
H^{\mathrm{Re}}_{GGPP} &= (H^{\mathrm{Re}}_{GGPP}(v,l) : v = 0,\ldots,4; l = 1,\ldots,L)\\
H^{\mathrm{Im}}_{GGPP} &= (H^{\mathrm{Im}}_{GGPP}(v,l) : v = 0,\ldots,4; l = 1,\ldots,L)\\
H^{\mathrm{Re}}_{LGPP} &= (H^{\mathrm{Re}}_{LGPP}(u,v,l) : u = 0,\ldots,7; v = 0,\ldots,4\\
&\qquad l = 1,\ldots,L)\\
H^{\mathrm{Im}}_{LGPP} &= (H^{\mathrm{Im}}_{LGPP}(u,v,l) : u = 0,\ldots,7; v = 0,\ldots,4\\
&\qquad l = 1,\ldots,L) \qquad (17)
\end{aligned}$$

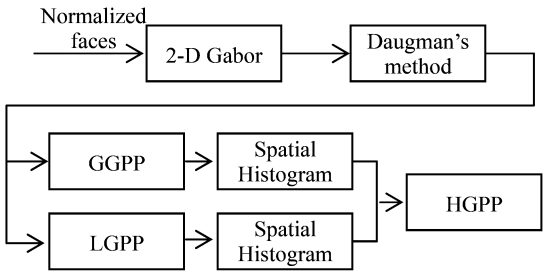

Fig. 8. Diagram of HGPP.

where L is the number of subregions divided for the histogram computation.

E. More Analysis of HGPP

So far, we have described the procedure to represent a face object by using HGPP as shown in Fig. 8. This section will discuss more about the advantages of the HGPP object representation.

First, HGPP is encoded from the Gabor features with multiple orientations and frequencies (scales). It is well known that Gabor wavelets can exhibit desirable characteristics of spatial locality, orientation selectivity, and spatial frequency (multiresolution). HGPP straightforwardly inherits these merits of Gabor transformation. Therefore, it has encoded plenty of discriminative information for object classification.

Second, it is very clear that the extraction of HGPP does not need any training procedure since it is directly derived from a single original image without using any external model. Thus, the generalizability problem is naturally avoided, compared with popular learning methods whose performance will be degraded if the distribution of the testing samples is quite different from that of the training set.

Third, the computation of HGPP is intrinsically efficient. One can easily find the fact that the computation procedure of GGPP and LGPP can be completed by only bit shift, and bit XOR operators. The local histogram computation can be achieved by add operator of integers ranged in from 0 to 255.

Finally, both the Gabor's localizability and the region-based histograms ensure that the final HGPP representation be a local model. Thus, it can be robust to local distortions, caused by accessory, expression variations, and so on.

III. Image Recognition Based on HGPP

As a kind of histogram-based object representation method, HGPP cannot be matched directly by the traditional distance measurements such as Euclidean distance. Although there are several methods for the histogram matching, such as histogram intersection, chi-square distance, in this paper, we mainly exploit histogram intersection as the similarity measurement.

In addition, it should be noted that HGPP is actually constituted by many pieces of histograms extracted from different spatial regions with different orientations and frequencies [refer to (16) and (17)]. It is obvious that these pieces of histograms contribute unequally to the discrimination of different objects. Therefore, we further present a weighted HGPP matching strategy by using the Fisher separation criterion.

A. HGPP Matching Based on Histogram Intersection

Given two histograms, H and S, their similarity based on histogram intersection is computed as follows:

$$S_{HI}(H,S)=\sum_{i=1}^{B}\min(H_i,S_i) \tag{18}$$

where B is the number of bins in the histogram and H_i denotes the frequency of GGPP or LGPP micro-patterns in the ith bin. Histogram intersection actually accumulates the common part of two histograms, and we formally formulate the similarity of two HGPPs, $H1$, $H2$, as shown in (19).

According to (16) and (17), let $H1$, $H2$ be $H1 = (H1^{\mathrm{Re}}_{GGPP}, H1^{\mathrm{Im}}_{GGPP}, H1^{\mathrm{Re}}_{LGPP}, H1^{\mathrm{Im}}_{LGPP})$ and $H2 = (H2^{\mathrm{Re}}_{GGPP}, H2^{\mathrm{Im}}_{GGPP}, H2^{\mathrm{Re}}_{LGPP}, H2^{\mathrm{Im}}_{LGPP})$, respectively. We define their similarity as

$$\begin{aligned}&S(H1,H2)\\&=S(H1^{\mathrm{Re}}_{GGPP},H2^{\mathrm{Re}}_{GGPP})+S(H1^{\mathrm{Im}}_{GGPP},H2^{\mathrm{Im}}_{GGPP})\\&\quad+S(H1^{\mathrm{Re}}_{LGPP},H2^{\mathrm{Re}}_{LGPP})+S(H1^{\mathrm{Im}}_{LGPP},H2^{\mathrm{Im}}_{LGPP})\end{aligned} \tag{19}$$

where the four items in (19) are computed as follows:

$$\begin{aligned}&S(H1^{\mathrm{Re}}_{GGPP},H2^{\mathrm{Re}}_{GGPP})\\&=\sum_{v=0}^{4}\sum_{l=1}^{L}S_{HI}(H1^{\mathrm{Re}}_{GGPP}(v,l),H2^{\mathrm{Re}}_{GGPP}(v,l))\end{aligned} \tag{20}$$

$$\begin{aligned}&S(H1^{\mathrm{Im}}_{GGPP},H2^{\mathrm{Im}}_{GGPP})\\&=\sum_{v=0}^{4}\sum_{l=1}^{L}S_{HI}(H1^{\mathrm{Im}}_{GGPP}(v,l),H2^{\mathrm{Im}}_{GGPP}(v,l))\end{aligned} \tag{21}$$

$$\begin{aligned}&S(H1^{\mathrm{Re}}_{LGPP},H2^{\mathrm{Re}}_{LGPP})\\&=\sum_{u=0}^{7}\sum_{v=0}^{4}\sum_{l=1}^{L}S_{HI}(H1^{\mathrm{Re}}_{LGPP}(u,v,l),H2^{\mathrm{Re}}_{LGPP}(u,v,l))\end{aligned} \tag{22}$$

$$\begin{aligned}&S(H1^{\mathrm{Im}}_{LGPP},H2^{\mathrm{Im}}_{LGPP})\\&=\sum_{u=0}^{7}\sum_{v=0}^{4}\sum_{l=1}^{L}S_{HI}(H1^{\mathrm{Im}}_{LGPP}(u,v,l),H2^{\mathrm{Im}}_{LGPP}(u,v,l))\end{aligned} \tag{23}$$

where L denotes the numbers of subregions for the histogram extraction.

The similarity measurement between two HGPPs is defined by (19). It is, in fact, the sum of similarities between the corresponding pieces of histograms in each HGPP.

B. Weighted HGPP Based on Fisher Separation Criterion (FSC)

Generally, different object areas are of different importance for object recognition. Therefore, it is easy to understand that the histogram pieces extracted from different regions with different Gabor parameters take unequal discriminative information. Different weights can be naturally set to different histogram pieces when matching two HGPPs. Formally, we can keep (19) unchanged while rewriting (20)–(23) as (24)–(27), shown at the bottom of the page, where the ω item in each equation denotes the weight for the corresponding piece of histogram, which should be prelearned in the training stage.

Setting appropriate weights for different histogram pieces is of great importance for the final performance of the object recognition system. We further present an approach to learn a suitable weight for each histogram piece based on the Fisher separation criterion (FSC) [30]. FSC is often used to evaluate the discriminative ability of the feature.

Given a training set containing C classes of objects, let the similarities of two histogram pieces from different samples of the same class compose the intraclass similarity space. Those samples from different classes compose the extra-class similarity space. Then, one can compute the FSC by the following formulations.

For the pth histogram piece, in HGPP, the mean and the variance, m_I^p and σ_I^p, of the intraclass similarities of H_p, can be computed by

$$m_I^p=\frac{1}{C}\sum_{i=1}^{C}\frac{2}{N_i(N_i-1)}\sum_{k=2}^{N_i}\sum_{j=1}^{k-1}S_{HI}\left(H_{i,j}^{P},H_{i,k}^{p}\right) \tag{28}$$

$$\sigma_I^p=\left(\frac{2}{\left(\sum_{i=1}^{C}N_i(N_i-1)\right)-2}\sum_{i=1}^{C}\sum_{k=2}^{N_i}\sum_{j=1}^{k-1}\left(S_{HI}\left(H_{i,j}^{P},H_{i,k}^{p}\right)-m_I^p\right)^2\right)^{\frac{1}{2}} \tag{29}$$

where $H_{i,j}^p$ denotes the H_p histogram piece in HGPP extracted from the jth sample of the ith class, and N_i is the sample number of the ith class in the training set. Similarly, the mean and the

$$S(H1^{\mathrm{Re}}_{GGPP},H2^{\mathrm{Re}}_{GGPP})=\sum_{v=0}^{4}\sum_{l=1}^{L}\omega^{\mathrm{Re}}_{G\mathrm{GPP}}(v,l)S_{HI}(H1^{\mathrm{Re}}_{GGPP}(v,l),H2^{\mathrm{Re}}_{GGPP}(v,l)) \tag{24}$$

$$S(H1^{\mathrm{Im}}_{GGPP},H2^{\mathrm{Im}}_{GGPP})=\sum_{v=0}^{4}\sum_{l=1}^{L}\omega^{\mathrm{Im}}_{GGPP}(v,l)S_{HI}(H1^{\mathrm{Im}}_{GGPP}(v,l),H2^{\mathrm{Im}}_{GGPP}(v,l)) \tag{25}$$

$$S(H1^{\mathrm{Re}}_{LGPP},H2^{\mathrm{Re}}_{LGPP})=\sum_{u=0}^{7}\sum_{v=0}^{4}\sum_{l=1}^{L}\omega^{\mathrm{Re}}_{L\mathrm{GPP}}(u,v,l)S_{HI}(H1^{\mathrm{Re}}_{LGPP}(u,v,l),H2^{\mathrm{Re}}_{LGPP}(u,v,l)) \tag{26}$$

$$S(H1^{\mathrm{Im}}_{LGPP},H2^{\mathrm{Im}}_{LGPP})=\sum_{u=0}^{7}\sum_{v=0}^{4}\sum_{l=1}^{L}\omega^{\mathrm{Im}}_{L\mathrm{GPP}}(u,v,l)S_{HI}(H1^{\mathrm{Im}}_{LGPP}(u,v,l),H2^{\mathrm{Im}}_{LGPP}(u,v,l)) \tag{27}$$

variance, m_E^p and σ_E^p, of the extra-class similarities can be computed by

$$m_E^p = \frac{2}{C(C-1)} \sum_{i=1}^{C-1} \sum_{j=i+1}^{C} \frac{1}{N_i N_j} \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} S_{HI}\left(H_{i,k}^p, H_{j,l}^p\right) \tag{30}$$

$$\sigma_E^p = \left(\frac{1}{\left(\sum_{i=1}^{C-1} \sum_{j=i+1}^{C} N_i N_j \right) - 1} \sum_{i=1}^{C-1} \sum_{j=i+1}^{C} \sum_{k=1}^{N_i} \sum_{l=1}^{N_j} \left(S_{HI}\left(H_{i,k}^P, H_{j,l}^p\right) - m_E^P \right)^2 \right)^{\frac{1}{2}}. \tag{31}$$

Finally, the weight for the H_p histogram piece, ω_p, is set by the following formulation:

$$\omega_p = \frac{\left(m_I^p - m_E^p\right)^2}{(\sigma_I^P)^2 + (\sigma_E^P)^2}. \tag{32}$$

ω_p is actually calculated in one dimension based on the Fisher criterion [30]. The above method for setting weights actually makes use of the discriminative power of each histogram piece, and larger weights are assigned to the histogram pieces with more discriminative information to separate different objects and vice versa.

It should be noted that the FSC-based weight assignment method breaks unexpectedly the nonlearning nature of HGPP, since one has to obtain a training set to learn the weights. To see the least, we also show experimentally that the performance of the basic HGPP matching (without weighting) has been very satisfactory.

IV. Experiments

To validate the proposed methods, we apply the proposed method to face recognition. Face recognition is one of the most challenging research topics in computer vision and patter recognition. Many approaches, such as Eigenface [31], Fisherface [32], Bayesian approach [33], local feature analysis (LFA) [34], Gabor Fisher classifier (GFC)[9], etc., are available for comparisons. LBP-based methods [22], [23] have been recently proposed and have achieved the state-of-the-art results on the FERET database. In this section, we have conducted comparison experiments on two large-scale face databases, FERET face database (with 1196 subjects) and CAS-PEAL-R1 face database (with 1040 subjects).

A. Experiments on Different Configurations of HGPP

In this part, we will show how some parameters of HGPP influence its final performance. These parameters include the size of the normalized image, the number of the subregions, and the number of the histogram bins. At the same time, we also check the different contributions of the real and imaginary parts, as well as those of LGPP and GGPP. These experiments are conducted on the FERET face database, which has been widely used to evaluate face recognition algorithms. We used the same Galley and Probe sets as the standard FERET evaluation protocol. Fa containing 1196 frontal images of 1196 subjects is used as Gallery, while Fb (1195 images of expression variations), Fc (194 images taking under different illumination conditions), Dup I (722 images taken later in time) and Dup II (234 images, a subset of Dup I), are the Probe sets. Dup II consists of images that were taken at least one year after the corresponding Gallery images. In our experiments, all faces are cropped according to the manually located eyes positions supplied with the FERET data, and the images are normalized to a given size (for example, $128 * 128$).

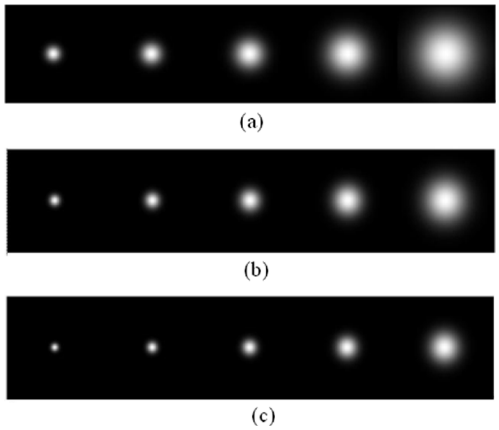

Fig. 9. Magnitude parts of the Gabor kernels at five different scales with different $f_{\max}$. (a) $f_{\max} = \pi/2$. (b) $f_{\max} = \sqrt{2}\pi/2$. (c) $f_{\max} = \pi$.

1) Experiment 1: On Image Size and Gabor Filters: Our first experiment aims to investigate the influence of the size of the normalized face image, as well as the frequency range of the Gabor filters (the parameter $f_{\max}$). In this experiment, we try to select a better $f_{\max}$ empirically. Three cases are considered, i.e., 128×128 image size with $f_{\max} = \pi/$ 2[7,8,9], 88×88 image size with $f_{\max} = \sqrt{2}\pi/2$, and 64×64 image size with $f_{\max} = \pi$. The idea behind this is that, as shown in Fig. 9, the size of the Gabor filters will be larger when $f_{\max}$ is decreased. Therefore, the images of large normalization sizes convoluting with the Gabor filters using (2) should be with a small $f_{\max}$ to capture the appropriate spatial frequency structure information. In [7], one can also find more discussions about the choice of $f_{\max}$. Previous works on the spatial histogram-based face recognition methods [22], [23] can guide us how to choose some appropriate parameters of HGPP. In this paper, GPP is empirically divided into 64 equal subregions to reserve more spatial information with 128 histogram bins per region to reduce the length of the histogram feature. The recognition results of the three cases are shown in Table I. We can easily find that HGPP using 128×128 image size with $f_{\max} = \pi/2$ is better than that using 88×88 image size with $f_{\max} = \sqrt{2}\pi/2$, and considerably outperforms that using 64×64 image size with $f_{\max} = \pi$. Experimental results show that the better performance can be achieved with the larger image size and an appropriate $f_{\max}$. Therefore, in the

TABLE I
RECOGNITION RATES FOR DIFFERENT SIZES OF THE NORMALIZED IMAGE (BASIC HGPP)

Size \ Probe	Fb	Fc	Dup I	Dup II
64x64, $f_{max}=\pi$	95.1	97.4	74.9	72.2
88x88, $f_{max}=\sqrt{2}\pi/2$	97.2	98.5	77.1	75.6
128x128, $f_{max}=\pi/2$	97.6	98.9	77.7	76.1

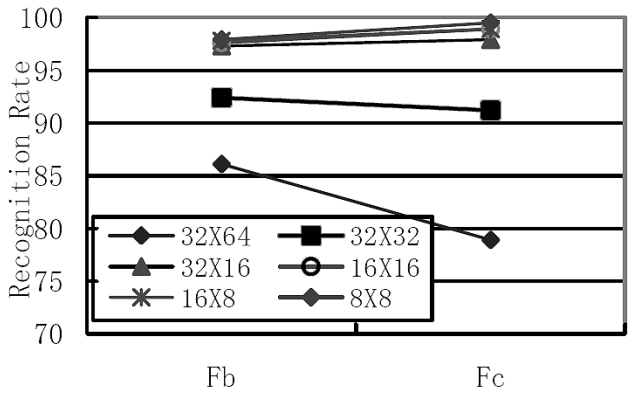

Fig. 10. Recognition rate for different numbers of subregions (basic HGPP).

following experiments, we fix the image size to be 128×128 with the Gabor parameter $f_{\max} = \pi/2$.

In experiments 2 and 3, two parameters to control the extraction of the face representation, the subregion size and the number of the histogram bins, are evaluated, which show us that HGPP is relatively robust and compact.

2) Experiment 2: On the Size of Subregion: After the input face image is transformed to the GPP domain, we use the subregion histograms (spatial histogram) to model the face. The advantage of the spatial histogram over the holistic histogram lies in its preservation of the spatial structure information. However, the size of the subregion has to be determined to balance the spatial locality and compactness of the model. Therefore, we do the following experiments on the Fb and Fc probe sets to examine the influence of the subregion size on the recognition rate. As the previous experiments suggest, we conduct the experiments based on the images of size 128×128 with $f_{\max} = \pi/2$. Six different subregion sizes, 32×64, 32×32, 32×16, 16×16, 16×8, and 8×8, are tested. As expected, a too-large subregion size may degrade the system due to the loss of much spatial information as shown in Fig. 10. However, a small subregion size will also result in the increase of the model complexity. For example, changing from the 8×8-sized subregion to the 16×16-sized subregion reduces the histogram feature length from 2949120 to 737280, while recognition rate is slightly decreased from 97.9% to 97.6% in the Fb probe set, and from 99.5% to 98.9% in the Fc probe set. Thus, a relative insensitivity to the subregion size can be observed in Fig. 10.

3) Experiment 3: On the Number of Histogram Bins: Our following experiments show that the performance of HGPP is also affected by the number of histogram bins. Using a small

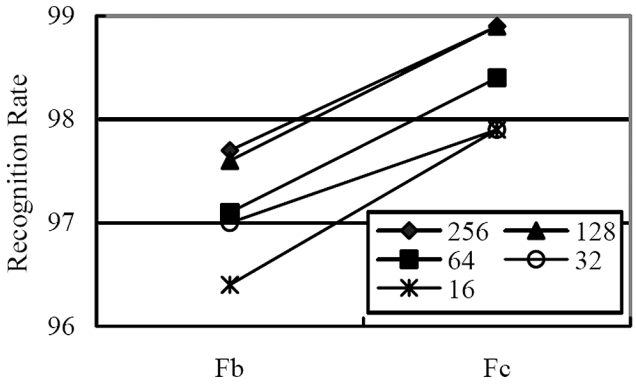

Fig. 11. Relationship between the numbers of histogram bins and the recognition rates (basic HGPP).

Fig. 12. Samples in the FERET database.

number of histogram bins makes the feature vector shorter but loses more information. However, the full-length histogram will also increase the model complexity and become more sensitive to the noise. One has to choose an appropriate number of bins in order to balance the efficiency and model complexity. In this paper, the uniform quantization method is used to partition the subregion histogram with equal intervals, i.e, $[0, \ldots, 256/B - 1], [256/B, \ldots, 2 * 256/B - 1], \ldots, [256 - 256/B, \ldots, 255]$, with B representing the number of histogram bins.

It is obvious that the length of the histogram feature will be greatly reduced when the number of histogram bins is changed from 256 to 16. However, the performance does not suffer a great deal. For example, changing from 256 bins to 16 bins, the recognition rate drops only from 97.9% to 96.4% on the Fb probe set, and from 98.9% to 97.7% on the Fc probe set as shown in Fig. 11.

4) Experiment 4: FSC Weights of Different Regions: $\omega_{L\mathrm{GPP}}(u, v, l)$ and $\omega_{G\mathrm{GPP}}(v, l)$ for the lth histogram piece are computed according to (32) using the 1002 frontal face images of 429 subjects in the FERET training database, and some samples can be seen in Fig. 12. In order to see the weight distributions, we visualized them as the images shown in Fig. 13. It should be noted that a dark intensity indicates a smaller weight; while a higher intensity means a bigger one. The samples shown in Fig. 13 are consistent with our intuition in that the eyes, nose, and mouth regions contribute more for

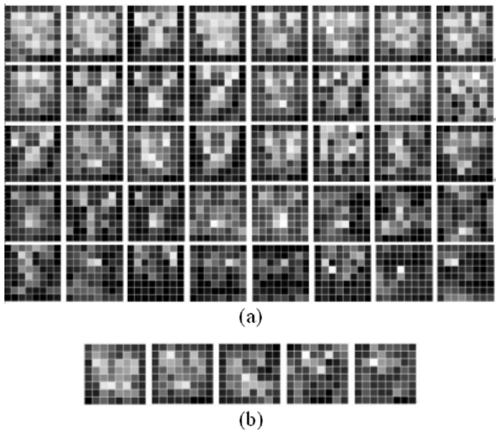

Fig. 13. (a) FSC weight for real part LGPP. (b) FSC weight for real part GGPP.

TABLE II
RANK-1 RECOGNITION RATES OF DIFFERENT GPPS ON THE FERET PROBES

Models \ Probes	Fb	Fc	Dup I	Dup II
Re_GGPP	91.9	87.6	64.3	59.4
Im_GGPP	91.2	86.1	64.4	61.1
Re_LGPP	97.0	97.9	74.9	72.6
Im_LGPP	96.8	97.9	74.7	74.4
Re_Im_GGPP	92.8	89.2	67.2	62
Re_Im_LGPP	97.1	98.5	77	75.6
Re_GGPP_LGPP	97.4	98.5	75.4	74.8
Im_GGPP_LGPP	97.0	98.5	76.1	75.2
GPP (all)	97.6	98.9	77.7	76.1

face recognition. Moreover, small scales with much rich texture information (as shown in Fig. 5) contain more discriminative information than large scales.

B. Comparisons of GGPP and LGPP of Real and Imaginary Parts

For the HGPP method, we actually have four different GPP models. These are GGPP of Gabor real part (hereinafter, Re_GGPP), GGPP of Gabor imaginary part (hereinafter, Im_GGPP), LGPP of Gabor real part (hereinafter, Re_LGPP), and LGPP of Gabor imaginary part (hereinafter, Im_LGPP). We make further experiments on Fb and Fc to evaluate these different GPP models. The results are given in Table II. From these results, we can safely conclude that the LGPP models (both Re_LGPP and Im_LGPP) generally outperform the GGPP models significantly. However, it should be noted that the GGPPs are less complex than the LGPPs, because the GGPP spatial histograms are extracted from only five GGPP maps. In comparison, there are 40 LGPP maps for the histogram extraction. This means that the space requirement of LGPP is 8 times that of GGPP.

From Table II, one can also find that the real and imaginary parts contribute similarly for face classification. And we further do experiments on their mutual compensations, as the results of Re_Im_GGPP and Re_Im_LGPP shown in Table II. One can see that by combining real and imaginary parts of Gabor coefficients, the performances of both GGPP and LGPP have been improved (though not very significantly). This means that the real and imaginary parts have provided different information for the classification.

In Table II, we also list the results of combining GGPP and LGPP models, as denoted by Re_GGPP_LGPP (combination of Re_GGPP and Re_LGPP) and Im_GGPP_LGPP (combination of Im_GGPP and Im_LGPP). The improvements are also observed compared with the original GGPP and LGPP. This indicates that GGPP and LGPP are also mutually compensational. In addition, all of information (both GGPP and LGPP of the real and imaginary parts) are combined together [denoted as GPP (all) in Table II]. Its rank-1 recognition rate is the highest, which has further verified the above observations.

TABLE III
RANK-1 RECOGNITION RATE COMPARISONS WITH OTHER STATE-OF-THE-ART RESULTS TESTED ON FERET PROBE SETS ACCORDING TO THE SAME FERET EVALUATION PROTOCOL

Methods	FERET probe sets			
	Fb	Fc	Dup I	Dup II
HGPP_Weighted	97.5	99.5	79.5	77.8
HGPP	97.6	98.9	77.7	76.1
LGBPHS_Weighted[41]	98.0	97.0	74.0	71.0
LGBPHS[41]	94.0	97.0	68.0	53.0
LBP_Weighted*[42]	97	79.0	66	64.0
LBP[42]	93.0	51.0	61.0	50.0
UMD LDA[21]	96.2	58.8	47.2	20.9
USC EBGM[21]	95.0	82.0	59.1	52.1
GFC	97.2	79.9	68.3	46.6

Fig. 14. Samples in the CAS-PEAL database.

C. Comparisons With Other State-of-the-Art Results Based on FERET Evaluation Protocol

To further validate the effectiveness of HGPP, we compare its performance with other known results reported on the four FERET probe sets according to the standard FERET evaluation protocol. So far, there have been several results available in the published literatures, such as the FERET'97 results published in 2000 [27], results of the LBP [23] published in 2004, and more recent results of the LGBPHS published in 2005 [22]. We compared our results with them, and the rank-1 recognition rates of

TABLE IV
EXPERIMENT RESULTS ON CAS-PEAL-R1 DATABASE (RANK-1 RECOGNITION RATE)

Probe sets \ Methods	Eigenface[28]	Fisherface[28]	GFC[28]	LGBPHS	HGPP	HGPP_Weighted
Accessory	37.1	61.0	85.1	86.8	91.9	92.5
Lighting	8.2	21.8	44.3	51	61.7	62.9
Expression	53.7	71.3	92.9	95.2	96.4	96.8
Distance	74.2	93.5	1.00	98.9	99.6	99.6
Background	80.5	94.4	98.9	98.7	99.6	99.8
Aging	50	72.7	93.9	1	96.9	98.4

these methods are shown in Table III. From this table, we can see that the proposed HGPP methods outperform all the other methods, especially on the Dup.I and Dup.II probe sets. Experiment results of these comparisons evidently illustrate that the proposed HGPP method (based on weighted matching) achieves the best results on the FERET face database. Since the face images in the FERET probe sets contain several sources of variation such as expression, lighting, and aging, these comparisons illustrate that HGPP is impressively robust to these extrinsic imaging conditions. Gabor features can exhibit the spatial frequency (scale), spatial locality and orientation selectivity properties corresponding to Gabor wavelets [9]. QBC, as a kind of quantification of Gabor feature, contributes to the robustness of HGPP. In [16], the robustness of QBC to the lighting variation has been verified. As a result, HGPP based on Gabor feature should be robust to the variation due to the illumination change.

The proposed method is also compared with the GFC method, which is based on the 64×64 normalized face images preprocessed by the histogram equalization. The augmented Gabor feature is down-sampled by a factor of 8 [9], and the dimensionality of the Gabor feature will be 2560. GFC is also trained on the FERET training database, and nearest-neighbor classifier based on the cosine similarity measurement is used to make a final classification [9]. HGPP has also achieved a better performance than GFC as shown in Table III.

Compared with the LGBPHS method presented by Zhang *et al.* in [22], the proposed HGPP method has made a further step on the state-of-the-art recognition rate on FERET face database. We must point out that HGPP in this paper is quite different from LGBPHS. In the LGBPHS method, Gabor magnitude (no phase information) is encoded by using LBP, while in this paper, we exploit the Gabor phase information by encoding the real and imaginary parts using QBC followed by LXP. This indicates that the phase part of the Gabor feature, which has been generally thought useless in face recognition, actually contains much discriminative information, and it can be well exploited for face classification using the proposed HGPP method.

D. Comparisons Based on CAS-PEAL-R1 Evaluation Protocol

More experiments are conducted on another large-scale face database, CAS-PEAL, for further validation of the proposed HGPP method. The CAS-PEAL face database was constructed under the sponsors of China National Hi-Tech Program [28]. Currently, the CAS-PEAL face database contains 99 594 images of 1040 individuals (595 males and 445 females) with varying pose, expression, accessory, and lighting (PEAL). The CAS-PEAL-R1, a subset of the CAS-PEAL face database, has been released for the purpose of research, which contains 9060 images of 1040 persons. Some samples are shown in Fig. 14. An accompanying evaluation protocol is provided, and the evaluation results of several well-known benchmarks including Eigenface, Fisherface, and GFC methods.

According to the CAS-PEAL-R1 evaluation protocol, experiments are conducted on the six CAS-PEAL-R1 probe sets, i.e., expression, accessory, lighting, background, distance, and aging. The training database contains 1200 images of 300 subjects [28]. The comparison results are listed in Table IV. It should be noted that the performances of benchmarks in the table are cited directly from [28]. Details about LGBPHS can refer to [22]. From the comparison results in Table IV, we can see that the HGPP method outperforms all the other methods, especially on the three *largest* probe sets, accessory, lighting, and expression. For instance, on the lighting probe set, the rank-1 recognition rate of our method is 62.9%, while that of the LGBPHS method is only 51%. The results illustrate that HGPP is much more robust than other methods against the extrinsic imaging conditions.

It should be noted that all parameters of the proposed basic HGPP method are the same as in the FERET database and NOT further tuned for the CAS-PEAL-R1 database. Therefore, we may conclude safely that the parameters of the proposed method can be easily controlled even for different databases.

V. SUMMARY AND FUTURE WORK

This paper proposes a novel object descriptor, named HGPP, which encodes the phase variation with the orientation changing of Gabor wavelet at a given scale, as well as the connection among local neighborhoods of the Gabor phase information. Objects are finally modeled as an ensemble of spatial histograms. The main contributions of the proposed method are as follows. 1) HGPP is not a learning-based face recognition method, therefore, the generalizability problem generally confronting the traditional learning-based methods is naturally avoided. 2) HGPP is quite different from the traditional Gabor usage in face recognition. In HGPP, only Gabor phase information is exploited to extract intrinsic object representation, while only magnitude information is exploited traditionally. 3) We have validated the HGPP method by conducting the experiments on two large-scale face databases, FERET and CAS-PEAL-R1, with more than 10 000 probes in total. On both databases, we have achieved better results than known published results. We should admit that face recognition is still an ongoing topic. However, we have made further step on the state-of-the-art recognition rate on the standard FERET database.

Analyzing the different parameters in the HGPP method, we noticed a relative insensitivity to the choice of region size and histogram bins. The choice of the frequency range seems to be related to the normalization image size. This is an interesting result since the currently dominant statistic or learning-based face modeling methods are more sensitive to their free parameters, and have to use an exhaustive training to find the optimal parameters.

Although high performance is achieved by the proposed method, we should also mention that one drawback of our method lies in the high-dimensional histogram features. There are several possible solutions to this problem. One straightforward way is to reduce the number of Gabor features (e.g., keeping only Im_LGPP or Re_LGPP) or adaptive binning method. Another one is using feature selection or discriminant analysis method to reduce the feature length. Due to its excellent performance, in face recognition, we also expect that the proposed method is applicable to other object recognition tasks as well.

ACKNOWLEDGMENT

The authors would like to thank B. Cao and W. Zhang for providing the results of the GFC and LGBPHS methods. They would also like to thank Prof. C. X. Ling from the University of Western Ontario and K. K. Williams from the Department of Biomedical Engineering, Georgia Institute of Technology, for helping us improve this paper, J. Chen for the fruitful discussion about the organization of this paper, and the anonymous reviewers for their kind suggestions.

REFERENCES

[1] A. Leonardis and H. Bischof, "Robust recognition using eigenimages," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 99–118, 2000.

[2] T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Comput. Vis. Image Understand.*, vol. 81, no. 3, pp. 231–268, 2001.

[3] B. Modhaddan and A. Pentland, "Probabilistic visual learning for object representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 696–710, Jul. 1997.

[4] T. Pajdla and J. Matas, "Human detection based on a probabilistic assembly of robust part detectors," in *Proc. 8th Eur. Conf. Computer Vision*, 2004, pp. 69–82.

[5] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. Roy. Soc. Lond. B*, vol. 207, pp. 187–217, 1980.

[6] D. Gabor, "Theory of communication," *J. Inst. Elect. Eng.*, vol. 93, no. 26, pt. III, pp. 429–457, 1946.

[7] M. Lades, J. C. Vorbruggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Wurtz, and W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Trans. Computers*, vol. 42, no. 3, pp. 300–311, Mar. 1993.

[8] L. Wiskott, J.-M. Fellous, N. Kruger, 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.

[9] 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.

[10] S. Shan, P. Yang, X. Chen, and W. Gao, "AdaBoost Gabor Fisher classifier for face recognition," in *Proc. IEEE Int. Workshop Analysis and Modeling of Faces and Gestures*, 2005, pp. 278–291.

[11] H. Szu, B. Telfer, and J. Garcia, "Wavelet transforms and neural networks for compression and recognition," *Neural Netw.*, vol. 9, no. 4, pp. 695–708, 1996.

[12] X. Chen, J. Yang, J. Zhang, and A. Waibel, "Automatic detection and recognition of signs from natural scenes," *IEEE Trans. Image Process.*, vol. 13, no. 1, pp. 87–99, Jan. 2004.

[13] R. Mehrotra, K. Namuduri, and N. Ranganathan, "Gabor filter-based edge detection," *Pattern Recognit.*, vol. 25, no. 12, pp. 1479–1494, 1992.

[14] J. Lee and S. D. Wang, "Fingerprint feature extraction using Gabor filters," *Electron. Lett.*, vol. 35, no. 4, pp. 288–290, 1999.

[15] J. G. 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.

[16] 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.

[17] M. J. Swain and D. H. Ballard, "Color indexing," *Int. J. Comput. Vis.*, vol. 7, no. 1, pp. 11–32, 1991.

[18] C. Schmid and R. Mohr, "Local greyvalue invariants for image retrieval," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 5, pp. 530–534, May 1997.

[19] M. Stricker and A. Dimai, "Color indexing with weak spatial constraints," in *Proc. SPIE, Storage and Retrieval for Still Image and Video Database IV*, 1996, vol. 2670, pp. 29–40.

[20] E. Hadjidemetriou, M. D. Grossberg, and S. K. Nayar, "Multiresolution histograms and their use for recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 7, pp. 831–847, Jul. 2004.

[21] S. Ravela and R. Manmatha, "Retrieving images by appearance," in *Proc. 6th IEEE Int. Conf. Computer Vision*, 1998, pp. 608–613.

[22] 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. 10th IEEE Int. Conf. Computer Vision*, 2005, pp. 786–791.

[23] T. Ahonen, A. Hadid, and M. Pietikainen, "Face recognition with local binary pattern," in *Proc. 8th Eur. Conf. Computer Vision*, 2004, pp. 469–481.

[24] H. Zhang, W. Gao, X. Chen, and D. Zhao, "Learning informative features for spatial histogram-based object detection," in *Proc. Int. Joint Conf. Neural Networks*, 2005, pp. 1806–1811.

[25] W. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, "Face recognition: A literature survey," *ACM Comput. Surv.*, pp. 399–458, 2003.

[26] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proc. IEEE*, vol. 83, no. 5, pp. 705–740, May 1995.

[27] 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.

[28] W. Gao, B. Cao, and S. Shan, "The CAS-PEAL Large-Scale Face Database and Evaluation Protocols," Tech. Rep. JDL_TR_04_FR_001, JDL, CAS, 2004.

[29] T. Ojala, M. Pietikainen, and T. Maenpaa, "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.

[30] R. M. Parry and I. Essa, "Feature weighting for segmentation," in *Proc. 5th Int. Conf. Music Information Retrieval*, 2004, pp. 116–119.

[31] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1991, pp. 586–590.

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

[33] B. Moghaddam, C. Nastar, and A. Pentland, "A Bayesian similarity measure for direct image matching," in *Proc. 13th Int. Conf. Pattern Recognition*, 1996, pp. 350–358.

[34] P. S. Penev and J. J. Atick, "Local feature analysis: A General statistical theory for object representation," *Network: Comput. Neural Syst.*, vol. 7, no. 3, pp. 477–500, 1996.

Baochang Zhang received the B.S. and M.S. degrees in computer science from the Harbin Institute of Technology, Harbin, China, in 1999 and 2001, respectively, where he is currently pursuing the Ph.D. degree.

His research interests include pattern recognition, machine learning, face recognition, and wavelets.

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

He is currently an Associate Researcher and serves as the Vice Director of the Digital Media Center, Institute of Computing Technology, CAS. He is also the Vice Director of the ICT-ISVision Joint R&D Lab for Face Recognition. His research interests cover image analysis, pattern recognition, and computer vision. He is especially focused on face recognition-related research topics.

Xilin Chen (M'00) 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 been a Professor at the Harbin Institute of Technology from 1999 to 2005. He was a Visiting Scholar at Carnegie Mellon University, Pittsburgh, PA, from 2001 to 2004. He jointed the Institute of Computing Technology, Chinese Academy of Sciences, Beijing, in August 2004. His research interests are image processing, pattern recognition, computer vision, and multimodal interface.

Dr. Chen has served as a program committee member for more than 20 international and national conferences. He has received several awards, including the China's State Scientific and Technological Progress Award in 2000, 2003, and 2005 for his research work.

Wen Gao (M'92–SM'05) received the B.Sc. and M.Sc. degrees in computer science from the Harbin University of Science and Technology, Harbin, China, and the Harbin Institute of Technology, Harbin, respectively, in 1982 and 1985, respectively, and the Ph.D. degree in electronics engineering from the University of Tokyo, Tokyo, Japan, in 1991.

He joined the Harbin Institute of Technology in 1985, where he served as a Lecturer, Professor, and Head of the Department of Computer Science until 1995. He was with Institute of Computing Technology, Chinese Academy of Sciences (CAS), Beijing, from 1996 to 2005. During his professor career with CAS, he was also appointed as Director of the Institute of Computing Technology, Executive Vice President of the Graduate School, as well as the Vice President of the University of Science and Technology of China. He is currently a Professor with the School of Electronics Engineering and Computer Science, Peking University, China. He has published four books and over 300 technical articles in refereed journals and proceedings in the areas of multimedia, video compression, face recognition, sign language recognition and synthesis, image retrieval, multimodal interface, and bioinformatics.

Dr. Gao is an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, Editor-in-Chief of the *Journal of Computer* (in Chinese), and Editor of the *Journal of Visual Communication and Image Representation*. He received the Chinese National Award for Science and Technology Achievement in 2000, 2002, 2003, and 2005.