Hindawi Publishing Corporation
EURASIP Journal on Advances in Signal Processing
Volume 2008, Article ID 469109, 8 pages
doi:10.1155/2008/469109

Research Article

Kernel Learning of Histogram of Local Gabor Phase Patterns for Face Recognition

Baochang Zhang,[1] Zongli Wang,[2] and Bineng Zhong[3]

[1] *School of Automation Science and Electrical Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100080, China*
[2] *Computer Science and Engineering, Beijing Institute of Technology, Beijing 100080, China*
[3] *Computer College, Harbin Institute of Technology, Harbin 150001, China*

Correspondence should be addressed to Baochang Zhang, bczhang@jdl.ac.cn

Received 27 August 2007; Revised 15 January 2008; Accepted 4 February 2008

Recommended by Hubert Cardot

This paper proposes a new face recognition method, named kernel learning of histogram of local Gabor phase pattern (K-HLGPP), which is based on Daugman's method for iris recognition and the local XOR pattern (LXP) operator. Unlike traditional Gabor usage exploiting the magnitude part in face recognition, we encode the Gabor phase information for face classification by the quadrant bit coding (QBC) method. Two schemes are proposed for face recognition. One is based on the nearest-neighbor classifier with chi-square as the similarity measurement, and the other makes kernel discriminant analysis for HLGPP (K-HLGPP) using histogram intersection and Gaussian-weighted chi-square kernels. The comparative experiments show that K-HLGPP achieves a higher recognition rate than other well-known face recognition systems on the large-scale standard FERET, FERET200, and CAS-PEAL-R1 databases.

1. INTRODUCTION

A good object representation or pattern representation is one of the key issues for a well-designed pattern recognition system. Representation issues include: what representation is desirable for the recognition of a pattern and how to effectively extract the representation from the original input signal. In face community, Gabor feature recently appears to be a promising way toward high accuracy face recognition. Gabor wavelet models quite well the receptive field profiles of cortical simple cells, therefore, Gabor feature can capture the salient visual properties such as the spatial localization, orientation selectivity, and spatial frequency characteristic [1]. Lades et al. [2] pioneer the use of Gabor wavelet for face recognition in the Dynamic Link Architecture framework. Wiskott et al. [3] subsequently develop 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, knows as a *jet*. Lyons et al. [4] have shown through experiments that the Gabor wavelet representation is optimal for classifying facial actions. The Gabor Fisher classifier (GFC) method proposed by Liu and Wechsler [5] is based on the magnitude part of Gabor feature, providing a promising way to enhance the face recognition performance. There are also some important applications of Gabor wavelet in sign recognition [6] and fingerprint recognition [7, 8]. It is easy for us to know that Gabor-based face recognition methods are mostly based on the magnitude part of Gabor feature. In fact, Gabor phase is very discriminative, and has been successfully used in iris and palm print identifications [9, 10].

Recently, Ahonen et al. [11] present a new approach based on local binary pattern (LBP) histograms for face recognition, considering both shape and texture information to represent the face images. Zhang et al. [12] combine the magnitude part of Gabor feature and the LBP operator, the so-called local Gabor binary pattern histogram sequence (LGBPHS) method, and achieved an excellent performance on the standard FERET database. Our former work, the so-called histogram of Gabor phase pattern (HGPP), encodes the Gabor phase variation derived from orientation change and local phase variations [13]. These methods are, in nature,

based on spatial histograms, which can capture the structure information of the input face object and provide an easy matching strategy.

In this paper, we propose a new kind of local Gabor phase pattern (LGPP) [13], from which local histograms are extracted and concatenated into a single extended histogram feature to capture the spatial information, named HLGPP. The recognition can be performed using the nearest-neighbor classifier with chi-square or histogram intersection as the similarity measurement. Moreover, histogram intersection (HI) [14] and Gaussian-weighted chi-squared (GW-chi) [15] functions have been proved to be positive definite, which were smoothly used in support vector machine (SVM) classifier [14, 15]. They show us that kernel methods can be successfully combined with the histogram feature, and motivate us to make kernel Fisher discriminant analysis for HLGPP (K-HLGPP). Experiments on the large-scale standard FERET, FERET200 [16], and CAS-PEAL [17] databases are performed to evaluate the effectiveness of HLGPP and K-HLGPP methods. Experimental results show that the proposed methods are much better than other well-known systems.

The rest of the paper is organized as follows. In Section 2, the background about the proposed method is introduced. In Section 3, HLGPP is proposed to extract the face representation from the original image. In Section 4, we propose a kernel learning method for HLGPP. In Section 5, experiments on the large-scale FERET, CAS-PEAL-R1, and FERET200 databases are conducted to evaluate the performances of the proposed methods. In the last section, some brief conclusions are drawn with some discussion on the future work.

2. BACKGROUND

Face Recognition is still an ongoing topic in computer vision research [18], because the current systems only perform well under the controlled environment but tend to fail in the complex situations with variations in different factors such as pose, illumination, expression, and so forth. Major approaches for face recognition in recent years are Eigenface [19], Fisherface [20], Bayesian method [21], Elastic Bunch Graph Matching (EBGM) [3], LBP-based methods [11, 12], and so forth. The performances of popular statistical or learning methods degrade abruptly, if the distribution of the testing samples is very different from that of the training set. Eigenface and Fisherface are the statistic methods based on principal component analysis (PCA) and Fisher discriminant analysis (FDA), which are linear feature extraction approaches. The Bayesian method uses a probabilistic measure of similarity to divide intensity difference into *extrapersonal* and *intrapersonal* spaces. In recent years, the kernelized feature extraction methods have been paid much attention, such as kernel principal component analysis (KPCA) [22] and kernel Fisher discriminant analysis (KFDA) [23, 24], which are nonlinear extensions to PCA and FDA, respectively. The selection of kernel function is one of open problems for the kernel-based methods, and some simple mercer's kernels are available, such as polynomial, Gaussian, RBF, and so on. We also find that some special kernel functions, GW-chi [15] and HI-kernel [14], have been successfully used in the field of computer vision. In this paper, we use the histogram-based HI and GW-chi kernel functions to make discriminant analysis for HLGPP.

2.1. Kernel Fisher discriminant analysis

The idea of KFDA is to yield a nonlinear discriminant analysis in a higher dimensional space. The input data is first projected into an implicit feature space F by the nonlinear mapping $\Phi : x \in R^N -> f \in F$, and then seek to find a transformation, maximizing the between-class scatter and minimizing the within-class scatter in F [25]. In its implementation, Φ is implicit and we just compute the inner product of two vectors in F by using a kernel function:

$$k(x, y) = (\Phi(x) \cdot \Phi(y)). \tag{1}$$

The between-class scatter matrix $\mathbf{S}_b$ and within-class scatter matrix $\mathbf{S}_w$ in the feature space F are defined as follows:

$$\begin{aligned} \mathbf{S}_b &= \sum_{i=1}^{C} p(\varpi_i)(u_i - u)(u_i - u)^T, \\ \mathbf{S}_W &= \sum_{i=1}^{C} p(\varpi_i) E\left\{ \left((\Phi(x_i) - u_i)(\Phi(x_i) - u_i)^T \right) \mid \varpi_i \right\}, \end{aligned} \tag{2}$$

$u_i = (1/n_i)\sum_{j=1}^{n_i} \phi(x_{ij})$ denotes the sample mean of class i, u is the mean of all training images in F, and $p(\varpi_i)$ is the prior probability. To perform FDA in F, it is equal to maximize (3).

$$J(w) = \frac{\mathrm{tr}(\mathbf{S}_b)}{\mathrm{tr}(\mathbf{S}_w)}. \tag{3}$$

Because any solution $\mathbf{w} \in F$ should lie in the span of all the samples in F, there exists

$$\mathbf{w} = \sum_{i=1}^{n} \alpha_i \phi(x_i). \tag{4}$$

Then we get the following maximizing criterion:

$$J(\boldsymbol{\alpha}) = \frac{\boldsymbol{\alpha}^T \mathbf{K}_b \boldsymbol{\alpha}}{\boldsymbol{\alpha}^T \mathbf{K}_w \boldsymbol{\alpha}}, \tag{5}$$

where $\mathbf{K}_w$ and $\mathbf{K}_b$ are defined as follows:

$$\begin{aligned} \mathbf{K}_w &= \sum_{i=1}^{C} p(\varpi_i) E(\eta_j - m_i)(\eta_j - m_i)^T, \\ \mathbf{K}_b &= \sum_{i=1}^{C} p(\varpi_i)(m_i - \overline{m})(m_i - \overline{m})^T, \end{aligned} \tag{6}$$

where $\eta_j = (k(x_1, x_j), k(x_2, x_j), \ldots, k(x_n, x_j))^T$, $m_i = ((1/n_i) \times \sum_{j=1}^{n_i} k(x_1, x_j), (1/n_i)\sum_{j=1}^{n_i} k(x_2, x_j), \ldots, (1/n_i)\sum_{j=1}^{n_i} k(x_n, x_j))^T$, and $\bar{m}$ is the mean of all η_j.

This problem can be solved by finding the leading eigenvectors of $\mathbf{K}_w^{-1}\mathbf{K}_b$, the so-called generalized kernel Fisher discriminant (GKFD) criterion. In our paper, we

use the technique of the pseudoinverse of the within-class scatter matrix, and then perform PCA on $\mathbf{K}_w^{-1}\mathbf{K}_b$ to get the transformation matrix $\boldsymbol{\alpha}$. The projection of a data point x onto $\mathbf{w}$ in F is given by:

$$v = (\mathbf{w}.\Phi(x)) = \sum_{i=1}^{n} \boldsymbol{\alpha}_i k(x_i, x). \tag{7}$$

In (1), if the x, y is the histogram feature, the kernel function can be redefined as follows:

$$\begin{aligned} k(x,y) &= K_{\mathrm{HI}}(x,y), \qquad k(x,y) = K_{\text{GW-chi}}(x,y), \\ K_{\mathrm{HI}}(x,y) &= S_{\mathrm{HI}}(x,y) = \sum_{i=1}^{B} \min(x_i, y_i), \end{aligned} \tag{8}$$

where $S_{\mathrm{HI}}(x, y)$ is histogram intersection, which actually accumulates the common parts of two histograms.

$$K_{\text{GW-chi}}(x,y) = \exp\left(-r^* S_{\text{GW-chi}}(x,y)\right), \tag{9}$$

where $S_{\text{GW-chi}}(x, y)$ is the chi-square statistic, B is the number of bins in the histogram, r is a constant, and x_i, y_i denote the frequency.

2.2. Daugman's method

Gabor wavelets (kernels, filters) can be defined as:

$$\psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} e^{(-\|k_{u,v}\|^2 \|z\|^2/2\sigma^2)} \left[e^{ik_{u,v}z} - e^{-\sigma^2/2} \right], \tag{10}$$

where $\overrightarrow{k_{u,v}} = \begin{pmatrix} k_{jx} \\ k_{jy} \end{pmatrix} = \begin{pmatrix} k_v \cos\phi_u \\ k_v \sin\phi_u \end{pmatrix}$, $k_v = f_{\max}/2^{v/2}$, $\phi_u = u(\pi/8)$, $v = 0, \ldots, 4, u = 0, \ldots, 7$, v is the frequency, and u is the orientation, with $f_{\max} = \sqrt{2}\pi$. For a given image z, the Gabor wavelet transformation can be defined as:

$$G_{u,v}(z) = I(z) * \Psi_{u,v}(z), \tag{11}$$

where $z = (x, y)$, $*$ denotes the convolution operator, and $G_{u,v}(z)$ is the convolution result corresponding to the Gabor kernel at scale $v = 0, \ldots, 4$ and orientation $u = 0, \ldots, 7$. It is well known that the magnitude part varies slowly with the spatial position, while the phases rotate in some rate with position. However, 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, and each pixel in the resultant image is encoded to two bits, $(P^{Re}_{u,v}(Z), P^{\mathrm{Im}}_{u,v}(Z))$, by the following rules:

$$\begin{aligned} P^{Re}_{u,v}(Z) &= \begin{cases} 0, & \text{if } \mathrm{Re}(G_{u,v}(Z)) > 0, \\ 1, & \text{if } \mathrm{Re}(G_{u,v}(Z)) \le 0, \end{cases} \\ P^{\mathrm{Im}}_{u,v}(Z) &= \begin{cases} 0, & \text{if } \mathrm{Im}(G_{u,v}(Z)) > 0, \\ 1, & \text{if } \mathrm{Im}(G_{u,v}(Z)) \le 0, \end{cases} \end{aligned} \tag{12}$$

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

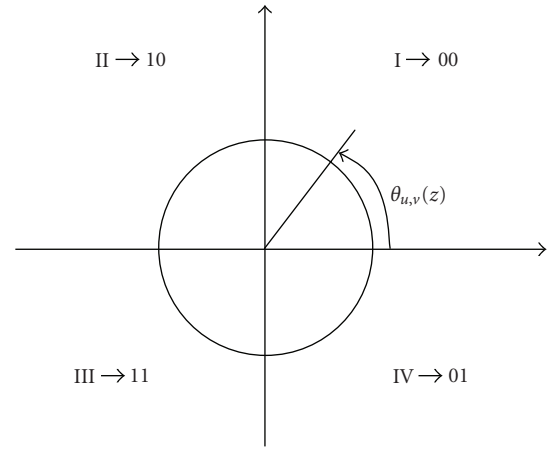

Figure 1: Quadrant bit coding.

3. HLGPP: AN OBJECT REPRESENTATION APPROACH

In this section, we propose a new kind of LGPP, which encodes the local neighborhood variations of Gabor phase at each orientation and scale. And LGPPs are combined with the local histograms to model the original face.

3.1. Quadrant bit coding (QBC) of Gabor phase angle

As shown in Figure 1, (12) can be reformulated as:

$$\begin{aligned} P^{Re}_{u,v}(Z) &= \begin{cases} 0, & \text{if } \theta_{u,v}(Z) \in \{\mathrm{I}, \mathrm{IV}\}, \\ 1, & \text{if } \theta_{u,v}(Z) \in \{\mathrm{II}, \mathrm{III}\}, \end{cases} \\ P^{\mathrm{Im}}_{u,v}(Z) &= \begin{cases} 0, & \text{if } \theta_{u,v}(Z) \in \{\mathrm{I}, \mathrm{II}\}, \\ 1, & \text{if } \theta_{u,v}(Z) \in \{\mathrm{III}, \mathrm{IV}\}. \end{cases} \end{aligned} \tag{13}$$

Thus, another bit code can be further obtained as follows:

$$P^{\mathrm{Atan}}_{u,v}(Z) = \begin{cases} 0, & \text{if } \theta_{u,v}(Z) \in \{\mathrm{I}, \mathrm{III}\}, \\ 1, & \text{if } \theta_{u,v}(Z) \in \{\mathrm{II}, \mathrm{IV}\}. \end{cases} \tag{14}$$

Specially, (14) reveals the relationship between the real and imaginary parts of Gabor feature. It is actually the XOR result of Daugman's two bit codes:

$$P^{\mathrm{Atan}}_{u,v}(Z) = P^{Re}_{u,v} \,\mathrm{XOR}\, P^{\mathrm{Im}}_{u,v}. \tag{15}$$

We call these three bit codes $P^{Re}_{u,v}$, $P^{\mathrm{Im}}_{u,v}$, $P^{\mathrm{Atan}}_{u,v}$ as quadrant bit coding (QBC) of the phase angle, since they are obtained according to the quadrants in which the phase angle lies.

3.2. Local Gabor phase pattern based on the local XOR pattern (LXP) operator

In this section, we propose to encode the local phase variations for each pixel with its neighborhood positions, the so-called LGPP. Formally, for each orientation u and

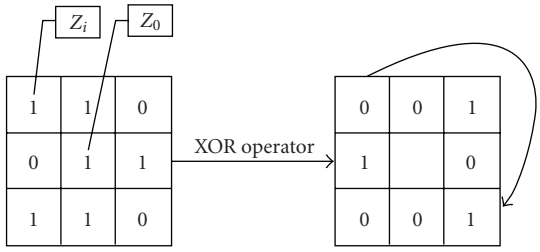

Figure 2: $\mathrm{LGPP}_{u,v}(Z_0)$ is a binary string 00101001.

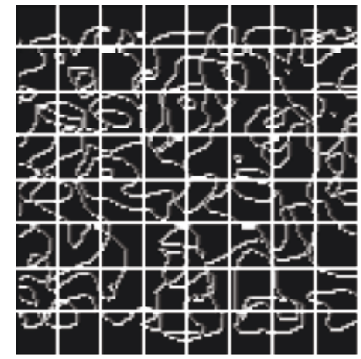

Figure 3: A sample of LGPP divided into 64 subregions.

frequency v, the real-, imaginary-, and atan-LGPP value at each pixel position are formulated as:

$$\begin{aligned}
\mathrm{LGPP}^{Re}_{u,v}(Z_0) &= \Big[P^{Re}_{u,v}(Z_0)\mathrm{XOR}\,P^{Re}_{u,v}(Z_1), P^{Re}_{u,v}(Z_0)\mathrm{XOR}\,P^{Re}_{u,v} \\
&\quad \times (Z_2),\ldots,P^{Re}_{u,v}(Z_0)\mathrm{XOR}\,P^{Re}_{u,v}(Z_8)\Big], \\
\mathrm{LGPP}^{\mathrm{Im}}_{u,v}(Z_0) &= \Big[P^{\mathrm{Im}}_{u,v}(Z_0)\mathrm{XOR}\,P^{\mathrm{Im}}_{u,v}(Z_1), P^{\mathrm{Im}}_{u,v}(Z_0)\mathrm{XOR}\,P^{\mathrm{Im}}_{u,v} \\
&\quad \times (Z_2),\ldots,P^{\mathrm{Im}}_{u,v}(Z_0)\mathrm{XOR}\,P^{\mathrm{Im}}_{u,v}(Z_8)\Big], \\
\mathrm{LGPP}^{\mathrm{Atan}}_{u,v}(Z_0) &= \Big[P^{\mathrm{Atan}}_{u,v}(Z_0)\mathrm{XOR}\,P^{\mathrm{Atan}}_{u,v}(Z_1), P^{\mathrm{Atan}}_{u,v}(Z_0)\mathrm{XOR} \\
&\quad \times P^{\mathrm{Atan}}_{u,v}(Z_2),\ldots,P^{\mathrm{Atan}}_{u,v}(Z_0)\mathrm{XOR}\,P^{\mathrm{Atan}}_{u,v}(Z_8)\Big],
\end{aligned} \tag{16}$$

where Z_i, $i = 1, 2, \ldots, 8$, is the 8-neighbors around the pixel position Z_0, and XOR denotes the bit *exclusive* or operator, the so-called local XOR pattern (LXP) operator [13] as shown in Figure 2. Eight neighbors can provide 8 bits to form a byte for each pixel, therefore, a decimal number ranged in $[0, 255]$ can be computed. Each value represents a mode how the Z_0 pixel is different from its neighbors.

By recalling the definition of QBC (16), the computation of each bit in (17) is actually equivalent to:

$$\begin{aligned}
&P^{Re}_{u,v}(Z_0)\mathrm{XOR}\,P^{Re}_{u,v}(Z_{\mathrm{i}}) \\
&\quad = \begin{cases} 0, & \text{if } \mathrm{Re}(G_{u,v}(Z_0)) \times Re(G_{u,v}(Z_i)) > 0, \\ 1, & \text{if } \mathrm{Re}(G_{u,v}(Z_0)) \times Re(G_{u,v}(Z_i)) \le 0, \end{cases} \\
&P^{\mathrm{Im}}_{u,v}(Z_0)\mathrm{XOR}\,P^{\mathrm{Im}}_{u,v}(Z_1) \\
&\quad = \begin{cases} 0, & \text{if } \mathrm{Im}(G_{u,v}(Z_0)) \times \mathrm{Im}(G_{u,v}(Z_i)) > 0, \\ 1, & \text{if } \mathrm{Im}(G_{u,v}(Z_0)) \times \mathrm{Im}(G_{u,v}(Z_i)) \le 0, \end{cases} \\
&P^{\mathrm{Atan}}_{u,v}(Z_0)\mathrm{XOR}\,P^{\mathrm{Atan}}_{u,v}(Z_{\mathrm{i}}) \\
&\quad = \begin{cases} 0, & \begin{array}{l} \text{if}\Big(Re(G_{u,v}(Z_0)) \times \mathrm{Im}(G_{u,v}(Z_0))\Big) \\ \times\Big(\mathrm{Re}(G_{u,v}(Z_i)) \times \mathrm{Im}(G_{u,v}(Z_i))\Big) > 0, \end{array} \\ 1, & \begin{array}{l} \text{if}\Big(Re(G_{u,v}(Z_0)) \times \mathrm{Im}\,(G_{u,v}(Z_0))\Big) \\ \times\Big(\mathrm{Re}(G_{u,v}(Z_i)) \times \mathrm{Im}(G_{u,v}(Z_i))\Big) \le 0. \end{array} \end{cases}
\end{aligned} \tag{17}$$

From (17), one can clearly know that LGPP actually encodes the sign difference of the central pixel from its neighbors, or reveals the relationships between neighbors whether they are in the same quadrants.

3.3. Histogram of local Gabor phase pattern

In Daugman's iris recognition method, quadrant-bit codes are directly used to form the representation of an iris image, and classification is achieved by the hamming distance. To model LGPPs more efficiently and compactly, in this paper, we exploit the spatial histogram to represent the distribution of the encoded micropatterns.

However, a single global histogram suffers from losing the structure information of the object, and the spatial structure information is of the high importance for face recognition. In order to reserve the spatial information in the histogram features, LGPPs are spatially divided into nonoverlapping rectangular regions represented by $R_1, \ldots, R_L$, from which local histogram features are extracted, respectively (shown in Figure 3), and all 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, that is, JLHF of LGPP images, histogram of local Gabor phase pattern (HLGPP).

Formally, the HLGPP extraction procedure is formulated as:

$$\begin{aligned}
H_{\mathrm{LGPP}} = (&H_{\mathrm{LGPP}}(u, v, l) : u = 0, \ldots, 7; \\
&v = 0, \ldots, 4;\ l = 1, \ldots, L),
\end{aligned} \tag{18}$$

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

4. FACE RECOGNITION BASED ON HLGPP

As a kind of histogram-based object representation method, HLGPP cannot be matched effectively by the traditional distance measurements such as the Euclidean distance. There exist several methods for the histogram matching, such as histogram intersection, chi-square distance. In this paper, we mainly exploit the chi-square as the similarity measurement.

4.1. Direct HLGPP matching method

The chi-square distance is used to measure the similarity between two histograms, and we formally formulate the similarity of two HLGPPs, $H1, H2$, as follows:

$$\begin{aligned} &S^{u,v}_{\text{GW-chi}}(H1_{\text{LGPP}}, H2_{\text{LGPP}}) \\ &\quad = \sum_{l=1}^{L} S_{\text{GW-chi}}(H1_{\text{LGPP}}(\text{u},v,l), H2_{\text{LGPP}}(u,v,l)), \\ &S(H1_{\text{LGPP}}, H2_{\text{LGPP}}) \\ &\quad = \sum_{u=0}^{7}\sum_{v=0}^{4} S^{u,v}_{\text{GW-chi}}(H1_{\text{LGPPI}}, H2_{\text{LGPPI}}), \end{aligned} \tag{19}$$

where L denotes the number of subregions for histogram extraction.

In the traditional statistic-based face recognition method, a training procedure is often necessary to extract the face representation. The advantage of the leaning-based methods lies in that they can use the background information, such as the variations due to expression, lighting, and aging changes, contained in a given training dataset, which is often offered by the face recognition test protocol, that is FERET. In the following part, we present how HLGPP makes discriminant analysis based on the HI and GW-chi kernels, which show that it can be easily combined with the statistic or leaning-based methods.

4.2. Kernel learning for HLGPP (K-HLGPP)

In this section, the proposed spatial histogram based kernel Fisher discriminant analysis method is used to find a discriminant transformation space, which is a prelearning way to use the background information. Formally, for spatial histogram feature extracted from each local region, a transformation matrix $\mathbf{w}_i$ can be calculated by the kernel Fisher method with HI and GW-chi kernels shown in Section 2, and then v_i is the extracted feature calculated by using (20):

$$v_i = \mathbf{w}_i \Phi(x) = \sum_{j=1}^{n} \boldsymbol{\alpha}_i^j k(x_i^j, x), \tag{20}$$

x_i^j is the histogram feature for the local region R_i of the jth face image, and v^1, v^2 are the feature vectors corresponding to two face images P_1, P_2. The similarity rule based on the cosine similarity between the corresponding extracted feature vectors is defined as follows:

$$d(P_1, P_2) = \sum_{i=1}^{L} \frac{v_i^1 \cdot v_i^2}{||v_i^1|| \cdot ||v_i^2||}. \tag{21}$$

From (21), we can easily know that the proposed method is based on the sum rule. It can actually use the spatial structure information of the face image, therefore, it should be appropriate to face recognition.

Table 1: Rank-1 recognition rate for different HLGPPs.

Methods	Fb	Fc	Dup1	Dup2
LGBPHS	94.7	97.0	58.8	49.0
Re_HLGPP	95.1	96.9	70.5	69.6
Im_HLGPP	95.8	97.9	71.1	67.9
Atan_HLGPP	96.1	98.5	73.7	69.6
Atan_K-HLHPP$_{\text{HI}}$	97.3	98.9	74.2	68.4
Atan_K-HLGPP$_{\text{GW-chi}}$	97.99	99.5	77.9	72.6

Table 2: Recognition rates for different sizes of the subregion (direct Atan_HLGPP).

Subregion size	Probe sets	
	Fb (%)	Fc (%)
16×16	94.3	98.5
8×16	95.1	98.5
8×8	96.1	98.5
8×4	95.8	99.5

5. EXPERIMENTS

To compare the performances of the proposed method and other well-known face recognition methods, the experiments are conducted on the standard FERET, CAS-PEAL-R1, and FERET200 databases, respectively.

5.1. Experiments on the standard FERET database

We have tested the proposed method on the standard FERET database [16], which is widely used to evaluate the face recognition algorithms. In the experiments, all images are cropped to the size of 64×64 according to the manually located eye positions supplied with the FERET database. We use the same gallery and probe image sets as in the standard FERET test. Fa (1196 images for 1196 subjects) is the gallery database, while Fb (1195 images), Fc (194 images), Dup I (722 images), and Dup II (234 images) are used as the probe sets.

Experiment 1: on different HLGPPs

In this part, we evaluate the performances of the HLGPPs face representation based on three kinds of QBC schemes on all the probe sets of the standard FERET database, and 64 subregions for the 64×64 normalized face images are chosen to reserve more structure information.

From Table 1, we can see that Atan_HLGPP achieves a better performance than Re_HLGPP and Im_HLGPP, partly because QBC of Atan_HLGPP reveals the relationship between real and imaginary parts of Gabor feature, and Re_HLGPP or Im_HLGPP just consider the real or imaginary part Gabor feature. HLGPP gets a much better results than LGBPHS using the same parameters, which confirms that the proposed method can provide a more effective face representation. The GW-chi kernel ($r = 0.00005$) achieves

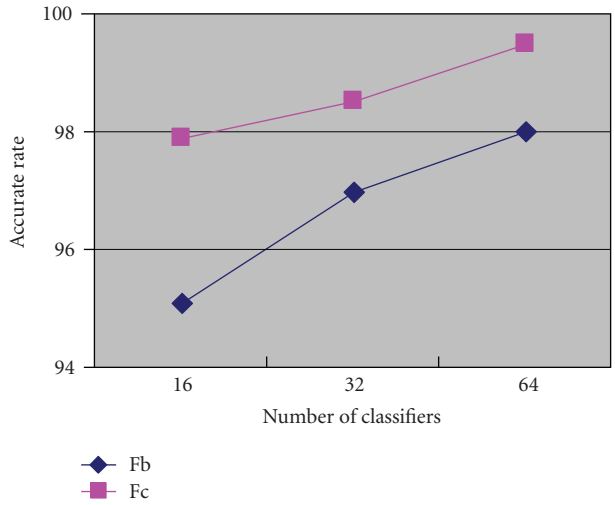

Figure 4: Performance of Atan_K-HLGPP for different number of classifiers on FERET Fb and Fc.

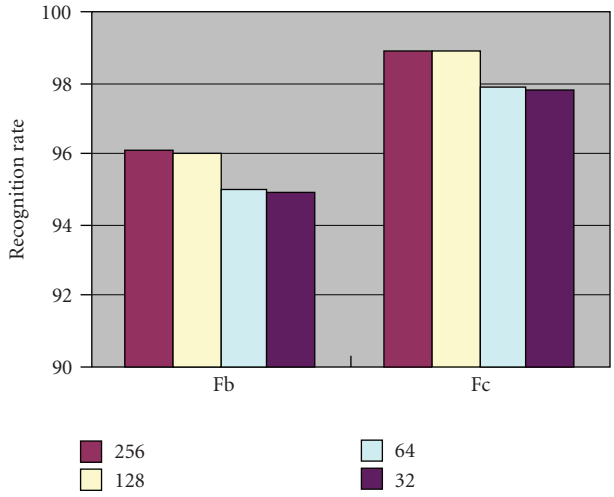

Figure 5: Relationship between the number of histogram bins and recognition rate (direct Atan_HGLXP).

a higher recognition rate than the HI-kernel, because it can capture the complex variations existed in a training database.

Experiment 2: on different subregion sizes

The advantage of the spatial histogram over holistic histogram lies in its preservation of the spatial information. We do the following experiments to examine the influence of the subregion size on the recognition rate on FERET-Fb and FERET-Fc. Four different subregion sizes, 16×16, 8×16, 8×8, 8×4, are tested. From Table 2, as expected, a too large subregion size degrades the system due to the loss of much spatial information for Atan_HLGPP. In Figure 4, we also evaluate the performance of K-HLGPP when different numbers of classifiers are used for the final classification, which shows that a larger number of classifiers result in a performance increase.

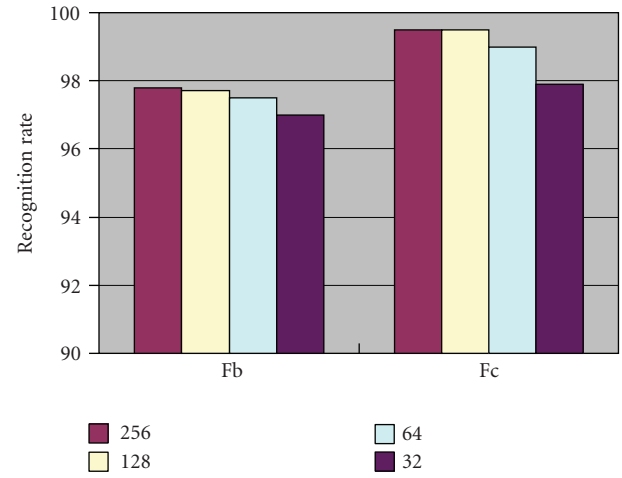

Figure 6: Relationship between the number of histogram bins and recognition rate (Atan_K-HGLXP).

Table 3: Rank-1 recognition rate comparisons with other state-of-the-art results tested on FERET probe sets according to the standard FERET evaluation protocol.

	Fb	Fc	Dup I	Dup II
K-HLGPP	**98.9**	**99.5**	**81**	**75.6**
Atan_K-HLGPP	**97.99**	**99 . 5**	**77.9**	72.6
Atan_HLGPP	96.1	98.5	73.7	69.6
HGPP	95.1	97.4	74.9	72.2
LGBPHS	94	97	68	53
LBP	93	51	61	50
GFC	97.2	79.9	68.3	46.6

Experiment 3: on different numbers of histogram bins

In this paper, the uniform quantization method is used to partition the subregion histogram with equal intervals, that is, [0, . . . , 256/B-1], [256/B, . . . , 2*256/B-1], . . . , [255-256/B, . . . , 255] with B representing the number of histogram bins. It is obvious that the length of the histogram feature is greatly reduced when the number of histogram bins is changed from 256 to 32 as shown in Figures 5 and 6, however, the performance does not suffer a lot.

Experiment 4: Comparisons with other well-known face recognition systems based on FERET evaluation protocol

To further validate the effectiveness of HLGPP-based methods, we compare their performances with other well-known results reported on the four FERET probe sets according to the standard FERET evaluation protocol. There are several results available in the published literatures, such as the FERET'97 results published in 2000 [16], results of LBP [11] published in ECCV2004, and more recent results of LGBPHS published in ICCV2005 [12]. We compared our results with them, and the rank-1 recognition rates of these methods are shown in Table 3. From this table, we can see

Table 4: Experiment result on CAS-PEAL-R1 database (rank-1 recognition rate).

	Eigenface	Fisherface	GFC	LGBPHS	Atan_HLGPP	HGPP	Atan_K-HLGPP
Accessory	37.1	61.0	85.1	86.8	91.2	91.9	92.8
Lighting	8.2	21.8	44.3	51.0	57.9	61.7	**70.1**
Expression	53.7	71.3	92.9	95.2	96.1	96.4	96.9

that K-HLGPP outperforms all the other results lies in that it can use the background information, such as the variations due to expression, lighting, and aging changes, contained in the training set provided by the standard FERET protocol [16]. Results of these comparisons evidently illustrate that K-HLGPP (including three kinds of QBCs) achieves the best results on the FERET face database. It should be noted that the numbers of Atan_K-HLGPP and K-HLGPP are 128 and 32 to reduce the feature length, respectively. HGPP is also based on the 64×64 normalized face images, with 64 subregions and 128 histogram bins. Note that K-HLGPP uses the GW-chi kernel.

5.2. Experiments based on the CAS-PEAL-R1 evaluation protocol

More experiments are conducted on another large-scale face database, CAS-PEAL, for further validation of the proposed method. Part of the CAS-PEAL face database, named CAS-PEAL-R1, has been released for research purpose, which contains 9060 images of 1040 subjects. An accompanying evaluation protocol is provided, as well as the evaluation results of several well-known benchmarks including Eigenface, Fisherface, and Gabor Fisher Classifier (GFC). Experiments are conducted on three *largest* CAS-PEAL-R1 probe sets, that is, expression, accessory, and lighting. The training database contains 1200 images of 300 subjects. From the comparison results in Table 4, we can see that the K-HLGPP method outperforms all the other benchmarks, for instance, the rank-1 recognition rate of our method is 70.1%, while that of GFC is only 44.3% on the lighting probe set.

5.3. Experiments based on the FERET200 database

A good face recognition system is expected to tolerate pose, expression, and illumination variations. The proposed algorithm is tested on FERET200. This set includes 1400 images of 200 individuals (each individual has 7 images) with moderate pose, expression, and illumination variations [16, 25]. The images are named by two character strings as "ba," "bj," "bk," "be," "bd," "bf," and "bg." In this experiment, we randomly select 100 people as the training set. The other 100 people are used to test the proposed method. The "ba" part is used as the gallery images, and other images are as the probe images. We repeat this procedure 10 times, and the mean recognition rate and variance are used evaluate the performances of comparative methods.

The complexity is evaluated in terms of time consuming for feature extraction, which is key part of all comparative methods. To calculate the final feature for each face image in HGPP, Atan_HLGPP and Atan_K-HLGPP, we need 232 ms, 163 ms, and 268 ms using a 3.2 G CPU, 2 G RAM PC. The performances of the comparative methods are evaluated in terms of the rank-1 recognition rate. As shown in Table 5, Atan_HLGPP achieves the best performance and gets about 12% improvement than other comparative methods. For Atan_HLGPP and HGPP, they achieve similar performances while Atan_HLGPP saves 69 ms per image.

Table 5: Experiment result on FERET200 (rank-1 recognition rate).

	HGPP	Atan_HLGPP	Atan_K-HLGPP
Mean recognition rate	81.91	81.85	**93.83**
Variance	0.816556	0.529444	0.760111

6. CONCLUSIONS AND FUTURE WORK

Unlike traditional Gabor usage exploitingonly the magnitude information in face recognition, this paper proposes to encode the Gabor phase angle for face classification by quadrant bit coding (QBC)and local XOR pattern (LXP) operator. After coding the Gabor phaseby QBC, we further use the LXP operator to encode the local phase variations of QBC, and spatial region-based histograms are exploited as the final representation of a given face image, that is, histogram of local Gabor phase pattern (HLGPP). Two schemes are proposed to solve the face recognition problem, one is based on nearest-neighbor classifier with the chi-square distance as the similarity measure, and another is based on kernel analysis for HLGPP (K-HLGPP) to extract discriminative features for the final classification, which can use the background information contained in the training set. Our experiments showthat the proposed methods are impressively better than other well-known face recognition methods on the standard FERET, FERET200, and CAS-PEAL-R1 databases, and they are robust enough against the extrinsic imaging conditions.

Although the high performance is achieved in our paper, some improvements are still possible. One drawback of our method lies in the feature length. One of the possible directions is to speed up the system by some kinds of dimensionality reduction methods, for example, making feature selection to choose the more discriminative patterns. Due to its excellent performance, we expect that the proposed method can be applicable to other object recognition as well.

ACKNOWLEDGMENTS

B. Zhang appreciates the support from the JDL Lab at Chinese Academy of Sciences. Thanks are due to Professor Charles X. Ling from University of Western Ontario, and

Heather Ford from Griffith University for helping us to improve the paper. Thanks are also given to Yu Su from the JDL Lab for providing the result of the GFC method, and Pengfei Shan from the Chinese University of Hong Kong for improving the efficiency of the proposed method.

REFERENCES

[1] J. G. Daugman, "Two-dimensional spectral analysis of cortical receptive field problems," *Vision Research*, vol. 20, no. 10, pp. 847–856, 1980.

[2] Lades, J. C. Vorbrueggen, J. Buhmann, et al., "Distortion invariant object recognition in the dynamic link architecture," *IEEE Transactions on Computers*, vol. 42, no. 3, pp. 300–311, 1993.

[3] L. Wiskott, J.-M. Fellous, N. Kuiger, and C. von der Malsburg, "Face recognition by elastic bunch graph matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 775–779, 1997.

[4] M. J. Lyons, J. Budynek, A. Plante, and S. Akamatsu, "Classifying facial attributes using a 2-D Gabor wavelet representation and discriminant analysis ," in *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition (AFGR '00)*, pp. 202–207, Grenoble, France, March 2000.

[5] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Transactions on Image Processing*, vol. 11, no. 4, pp. 467–476, 2002.

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

[7] A. K. Jain, S. Prabhakar, and L. Hong, "A multichannel approach to fingerprint classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 348–359, 1999.

[8] C.-J. Lee and S.-D. Wang, "Fingerprint feature extraction using Gabor filters," *Electronics Letters*, vol. 35, no. 4, pp. 288–290, 1999.

[9] J. G. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1148–1161, 1993.

[10] D. Zhang, W.-K. Kong, J. You, and M. Wong, "Online palmprint identification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1041–1050, 2003.

[11] T. Ahonen, A. Hadid, and M. Pietikäinen, "Face description with local binary patterns: application to face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, 2006.

[12] 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 *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV '05)*, vol. 1, pp. 786–791, Beijing, China, October 2005.

[13] B. Zhang, S. Shan, X. Chen, and W. Gao, "Histogram of Gabor phase patterns (HGPP): a novel object representation approach for face recognition," *IEEE Transactions on Image Processing*, vol. 16, no. 1, pp. 57–68, 2007.

[14] A. Barla, F. Odone, and A. Verri, "Histogram intersection kernel for image classification," in *Proceedings of IEEE International Conference on Image Processing (ICIP '03)*, vol. 3, pp. 513–516, Barcelona, Spain, September 2003.

[15] S. Belongie, C. Fowlkes, F. N. Chung, and J. Malik, "Spectral partitioning with indefinite kernels using the nystorm extensions," in *Proceedings of the 7th European Conference on Computer Vision (ECCV '02)*, pp. 531–542, Copenhagen, Denmark, May 2002.

[16] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000.

[17] W. Gao, B. Cao, S. Shan, et al., "The CAS-PEAL large-scale chinese face database and baseline evaluations," *IEEE Transactions on Systems Man, and Cybernetics, Part A*, vol. 38, no. 1, pp. 149–161, 2007.

[18] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: a literature survey," *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003.

[19] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '91)*, pp. 586–591, Maui, Hawaii, USA, June 1991.

[20] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[21] B. Moghaddam, C. Nastar, and A. Pentland, "A Bayesian similarity measure for direct image matching," in *Proceedings of the 13th International Conference on Pattern Recognition (ICPR '96)*, vol. 2, pp. 350–358, Vienna, Austria, August 1996.

[22] B. Schölkopf, A. Smola, and K.-R. Müller, "Nonlinear component analysis as a kernel eigenvalue problem," *Neural Computation*, vol. 10, no. 5, pp. 1299–1319, 1998.

[23] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, and K.-R. Müller, "Fisher discriminant analysis with kernels," in *Proceedings of the 9th IEEE Workshop on Neural Networks for Signal Processing (NNSP '99)*, pp. 41–48, Madison, Wis, USA, August 1999.

[24] G. Baudat and F. Anouar, "Generalized discriminant analysis using a kernel approach," *Neural Computation*, vol. 12, no. 10, pp. 2385–2404, 2000.

[25] B. Zhang, X. Chen, S. Shan, and W. Gao, "Nonlinear face recognition based on maximum average margin criterion," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 554–559, San Diego, Calif, USA, June 2005.