Multi-view Gender Classification based on Local Gabor Binary Mapping Pattern and Support Vector Machines

Bin Xia, He Sun and Bao-Liang Lu* *Senior Member, IEEE*

Abstract— This paper proposes a novel face representation approach, local Gabor binary mapping pattern (LGBMP), for multi-view gender classification. In this approach, a face image is first represented as a series of Gabor magnitude pictures (GMP) by applying multi-scale and multi-orientation Gabor filters. Each GMP is then encoded as a LGBP image where a uniform local binary pattern (LBP) operator is used. After that, each LGBP image is divided into non-overlapping rectangular regions, from which spatial histograms are extracted. Although an LGBP feature vector can be obtained by fitting together the regional histograms, it can not be employed in pattern classification due to its high dimension. We propose that each regional LGBP feature be mapped onto a one-dimensional subspace independently before they are concatenated as a whole feature vector. This is attractive since we reduce the feature dimension and also preserve the spatial information of LGBP image. Two ways have been proposed to map the regional LGBP feature in this paper. One is so-called LGBMP-LDA using linear discriminant analysis (LDA) for dimensionality reduction while the other is to project the regional LGBP feature onto the class center connecting line, namely, LGBMP-CCL. As a result, despite several decades of Gabor filters, the final feature dimension is even less than that of the feature extracted by using LBP directly on gray-scale images. The classification tasks in our work are performed by support vector machines (SVM). The experimental results on the CAS-PEAL face database indicate that the proposed approach achieves higher accuracy than the others such as SVMs+Gray-scale pixel, SVMs+Gabor and SVMs+LBP approach, more particularly, it has the lowest dimension of feature vector.

I. INTRODUCTION

Human faces provide a lot of useful information representing gender, age as well as identity. Gender classification based on facial images is one of the most challenging problems for computer vision researchers [1]. It is also attractive due to its underlying application such as visual surveillance and robot vision [17]. Similar to most pattern classification problems, how to extract features and which pattern classifier to use are the two key points for gender classification.

Several classifiers have been introduced to gender classification, such as KNN, multi-layer neural networks [14], RBF networks [15] and SVMs [1]. It is important to find out which classifier is best for gender classification. B.Moghaddam *et al.* [1] have compared the performance of SVM with other pattern classifiers including Fisher linear discriminant, KNN, RBF networks and large ensemble-RBF classifiers. According to their experimental results, SVM seems to be

Bin Xia, He Sun and Bao-Liang Lu are with the Department of Computer Science and Engineering, Shanghai Jiao Tong University 800 Dong Chuan Rd., Shanghai 200240, China. *Corresponding author (email: blu@cs.sjtu.edu.cn).

much superior to all the others. Therefore, we pick SVM for multi-view gender classification in this paper.

To achieve the goal of extracting features from a facial image, four kinds of methods have been used in gender classification. The most straightforward method is to use merely the gray-scale or color pixel [1]. This kind of methods is very simple but fails to be effective due to variations of pose, expression and illumination. Another kind of methods is to project the facial image into another subspace, which has a much lower dimension in terms of the feature vector. These methods include principal component analysis (PCA), independent component analysis (ICA) and linear discriminant analysis (LDA) [2]. But these methods are not that robust to variations of face pose[4]. The third kind of methods attempts to incorporate facial feature detection with wavelet transformation to extract local facial features for gender classification, such as Gabor filter features [5] and the analysis of facial wrinkles and shapes [3]. The last kind of methods extracts local texture information from facial images such as local binary pattern (LBP) proposed by Ojala *et al.* [7] and Lian [4] has exploited this method for gender classification.

Zhang *et al.* [9] followed LBP and proposed a face recognition approach called local Gabor binary pattern histogram sequence (LGBPHS), which combines local binary pattern histogram and the magnitude part of Gabor filter. This approach is robust to noise and local image transformations caused by variations of illumination, occlusion and pose [9]. Nevertheless this approach is based on non-statistical learning and the dimension of feature vector extracted by this approach is however hundreds of times as large as that of the original LBP features (more than 500,000). Therefore it can not be employed to supervised learning problems due to the memory and time limit as well as the curse of dimensionality. In this paper, we improve the work of [9] and propose a new type of feature, local Gabor binary mapping pattern (LGBMP), which is obtained by mapping the regional local Gabor binary pattern (LGBP) feature into a low-dimensional space. Briefly speaking, an LGBP feature is a vector obtained by concatenating histograms region by region. Additionally, We propose two ways for dimensionality reduction, the LGBMP linear discriminant analysis (LGBMP-LDA) and LGBMP center connecting line (LGBMP-CCL). The dimension of our new feature is even less than that of the LBP feature directly extracted from gray-scale images.

The overall process of the proposed feature extraction approach based on LGBMP is illustrated in Fig. 1. In this
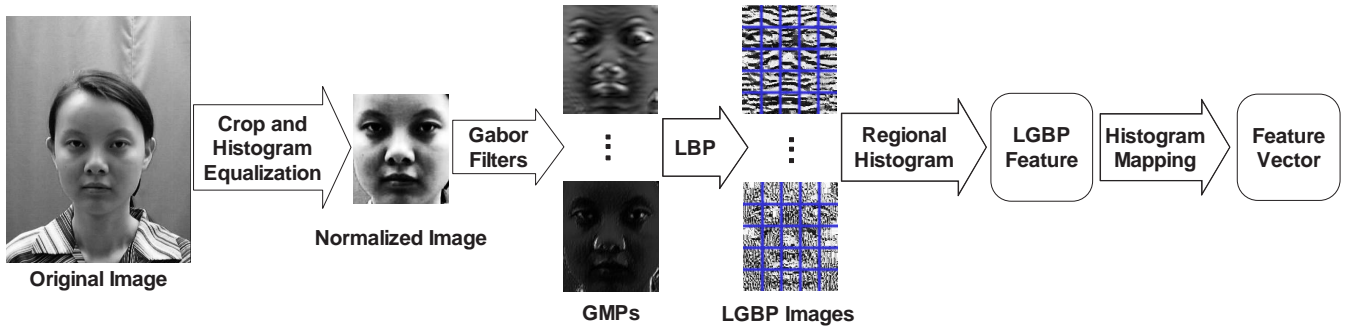
Fig. 1. The framework of the proposed LGBMP feature extraction method.

approach, a face image is modeled as a feature vector by the following procedure: (1) An input face image is pre-processed by locating eye positions, cropping, geometric normalization, and histogram equalization; (2) A series of Gabor magnitude pictures (GMP) are obtained by convolving the image with multi-scale and multi-orientation Gabor filters; (3) Each GMP is converted to an LGBP image by using $LBP_{8,1}^{u2}$ encoding method; (4) Each LGBP image is further divided into non-overlapping rectangular regions with specific size, and histograms are computed for all regions to form the LGBP feature. These steps are similar to [9]; (5) Each regional LGBP feature is mapped into a single value and all these values are concatenated to form the final feature vector. The two mapping solutions are shortly illustrated as follows: assume all the facial image in training set have been modeled as LGBP features. In all these features, the regional histograms with the same region order are gathered to form a new set. The number of new sets is the same as the total region number. We carry out dimensionality reduction on every such set independently. The LGBMP-LDA solution applies LDA to the set while LGBMP-CCL simply projects the whole set onto the center connecting line of the two class. Both of the two solutions map each histogram in the set into a single value. We substitute these generated values for their corresponding histograms in the original LGBP feature. We will discuss each step mentioned above in detail in the following sections. We also evaluate the proposed method on the CAS-PEAL face database [13] with comparison to SVMs+Gray-scale pixel, SVMs+Gabor and SVMs+LBP method. The experimental results indicate that the LGBMP methods outperform the others on both frontal and profile facial images.

The remaining part of the paper is organized as follows: Section II describes the computation of LGBP image in detail. Section III discusses the proposed dimensionality reduction approaches, LGBMP-LDA and LGBMP-CCL. And rich comparison experiments are presented in Section IV, including LGBMP-LDA versus LGBMP-CCL and LGBMP feature versus gray-scale, Gabor and LBP feature. Some brief conclusions are outlined in Section V with some discussions on future work.

II. FEATURE EXTRACTION WITH LOCAL GABOR BINARY PATTERNS

A. Multi-scale and multi-orientation Gabor filters

The LGBP feature is derived by combining Gabor magnitude part and uniform LBP histograms. Firstly, the multi-scale and multi-orientation Gabor filters are exploited to decomposite the normalized facial image. The Gabor filters are defined as follows [6]:

$$\Psi_{\mu,v}(z) = \frac{||k_{\mu,v}||^2}{\sigma^2} e^{(-\frac{||k_{\mu,v}||^2||z||^2}{2\sigma^2})} \left[e^{ik_{\mu,v}z} - e^{-\frac{\sigma^2}{2}} \right], \quad (1)$$

where μ and v define the orientation and the scale of the Gabor filters, $z = (x,y)^T$, $||\cdot||$ denotes the norm operator, and the wave vector is given by

$$\mathbf{k}_{\mu,v} = \left(\begin{array}{c} k_v \cos \Phi_\mu \\ k_v \sin \Phi_\mu \end{array} \right),$$

where $k_v = \frac{k_{max}}{\lambda^v}$ and $\Phi_\mu = \mu \frac{\pi}{8}$. k_{max} is the maximum frequency and λ is the spacing factor between wavelets in the frequency domain.

The Gabor representation of a facial image is the convolution of the image with a series of Gabor filters. Note that, only the magnitude part of the generated Gabor features is employed because it's not time-varying [9]. Let $I(x,y)$ be the facial image, its GMP with $\Psi_{\mu,v}(z)$ is defined as follows:

$$G(x,y,\mu,v) = I(x,y) * \Psi_{\mu,v}(z), \quad (2)$$

where $*$ denotes the convolution operator. In this paper, we use Gabor wavelets of five scales $v \in \{0,\ldots,4\}$ and eight orientations $\mu \in \{0,\ldots,7\}$ [10]. Therefore, 5×8=40 GMPs will be generated for each facial image.

B. Local Binary Patterns

According to [7], LBP has been proved a powerful way of texture description. The original LBP operator labels the pixels of an image by thresholding the 3×3-neighborhood of each pixel with the center value and considering the result as a binary number:

$$S(f_p - f_c) = \left\{ \begin{array}{ll} 1, & f_p \geq f_c \\ 0, & f_p < f_c, \end{array} \right. \quad (3)$$
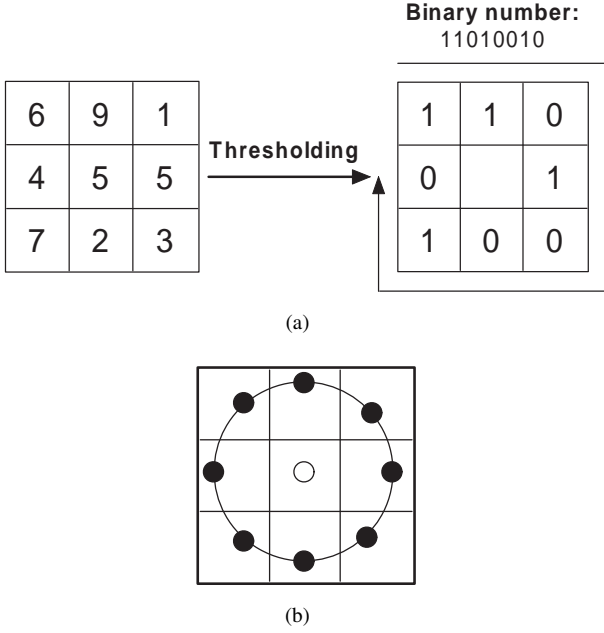
Binary number:
11010010

(a)

(b)

Fig. 2. (a) The basic LBP operator. (b) The circular $(8,1)$ neighborhood.

where f_c is the center value and $f_p (p = 0, 1, \cdots, 7)$ is the neighborhood of f_c. Then the LBP value at the center pixel is expressed as:

$$LBP(f_c) = \sum_{p=0}^{7} S(f_p - f_c)\, 2^p. \qquad (4)$$

A local binary pattern is called uniform if it contains at most two bitwise transitions from 0 to 1 or vice versa when the binary string is considered circular. For instance, 11111111, 11100011 are uniform patterns and 10101001 is not. We use the following notation for the LBP operator: $LBP_{P,R}^{u2}$ [8]. The subscript represents using the operator in a neighborhood of P sampling points on a circle of radius of R. The superscript $u2$ stands for using only uniform patterns and labeling all remaining non-uniform patterns with a single label. For $LBP_{P,R}^{u2}$, which is used in this paper, there are 58 labels for uniform patterns and one label for the remaining, thus 59 labels as a whole. Fig. 2 illustrates the basic LBP operator (Fig. 2 (a)) and an example of the $LBP_{8,1}^{u2}$ neighborhood (Fig. 2 (b)) where the pixel values are bilinearly interpolated whenever the sampling point is not in the center of a pixel.

C. Local Gabor Binary Patterns

An LGBP image can be obtained by applying the LBP operator to a GMP. Assume each LGBP image is divided into m non-overlapping rectangular regions $R_0, R_1, \ldots, R_{m-1}$. The L-bin histogram of j-th region in the v-scale and μ-orientation LGBP image is:

$$H_{\mu,v,j} = \{h_{\mu,v,0,j}, h_{\mu,v,1,j}, \ldots, h_{\mu,v,L-1,j}\}, \qquad (5)$$

where

$$h_{\mu,v,i,j} = \sum_{x,y} I\{f(x,y) = i\}\, I\{(x,y) \in R_j\},$$

and $f(x,y)$ is the LGBP image, $i = 0, 1, \ldots, L-1$, $j = 0, 1, \ldots, m-1$ and function I is:

$$I(A) = \begin{cases} 1, & \text{if } A \text{ is true} \\ 0, & \text{otherwise.} \end{cases} \qquad (6)$$

As mentioned in the previous subsection, we use $LBP_{8,1}^{u2}$ operator and thus each histogram has 59 bins ($L = 59$). Finally, histograms of all LGBP images are fitted together as the LGBP feature vector:

$$\begin{aligned} V = \{ & H_{0,0,0}, \ldots, H_{0,0,m-1}, \\ & H_{0,1,m-1}, \ldots, H_{7,4,m-1}\}. \end{aligned} \qquad (7)$$

III. DIMENSIONALITY REDUCTION

The dimension of the feature vector in formula (7) is 40 times more than that of the uniform LBP feature. It can be illustrated as follows: assume we divide each LGBP image into m regions, then the LBP feature has a dimension of $59m$ while the LGBP feature has $5\times8\times59m$=2,360m. If we set m=10$\times$10, then the total dimension is 236,000 which is a nightmare currently for many pattern classifiers. Several global methods can be used to reduce dimension such as PCA and randomly down-sampling [4]. However, these methods fail to take into account the spatial information as well as the label of samples. To cope with this problem, we reduce dimension independently on each region. The approach is presented as follows: suppose, in the training set, there are p images labeled as male and n images labeled as female. We rewrite formula (7) as:

$$V^{(k)} = \{H_{0,0,0}^{(k)}, \ldots, H_{\mu,v,j}^{(k)}, \ldots, H_{7,4,m-1}^{(k)}\}, \qquad (8)$$

where $V^{(k)}$ is the k-th feature vector in the training set and k is ranged from 1 to $p+n$. We also take into account the label information. Define $\omega^{(k)}$ as the label of the k-th sample:

$$\omega^{(k)} = \pm 1,$$

where +1 represents the male and -1 the female. Then we establish a new set $\mathcal{T}_{\mu,v,j}$ that takes all $H_{\mu,v,j}^{(k)}$ as its elements:

$$\mathcal{T}_{\mu,v,j} = \{H_{\mu,v,j}^{(k)}\}, \quad k = 1, 2, \ldots, p+n.$$

$\mathcal{T}_{\mu,v,j}$ can be further divided into two subsets according to the label:

$$\mathcal{T}_{\mu,v,j} = \mathcal{T}_{\mu,v,j}^{+} \bigcup \mathcal{T}_{\mu,v,j}^{-},$$

where

$$\begin{aligned} \mathcal{T}_{\mu,v,j}^{+} &= \{H_{\mu,v,j}^{(k)} | \omega^{(k)} = +1\} \\ \mathcal{T}_{\mu,v,j}^{-} &= \{H_{\mu,v,j}^{(k)} | \omega^{(k)} = -1\}. \end{aligned}$$

It is obvious that the size of set $\mathcal{T}_{\mu,v,j}^{+}$ is p and $\mathcal{T}_{\mu,v,j}^{-}$ is n. We compute the means for $\mathcal{T}_{\mu,v,j}^{+}$ and $\mathcal{T}_{\mu,v,j}^{-}$, or the two class

centers, as follows:

$$M_{\mu,v,j}^{+} = \frac{1}{p} \sum_{H_{\mu,v,j}^{(k)} \in \mathcal{T}_{\mu,v,j}^{+}} H_{\mu,v,j}^{(k)}$$

$$M_{\mu,v,j}^{-} = \frac{1}{n} \sum_{H_{\mu,v,j}^{(k)} \in \mathcal{T}_{\mu,v,j}^{-}} H_{\mu,v,j}^{(k)}. \qquad (9)$$

To map the histograms in each $\mathcal{T}_{\mu,v,j}$, we will present two solutions, LGBMP-LDA and LGBMP-CCL to find the project direction, notated as $D_{\mu,v,j}$, in the following subsections.

A. LGBMP-LDA

For each $\mathcal{T}_{\mu,v,j}$, we try to project the histograms in $\mathcal{T}_{\mu,v,j}$ onto a direction $D_{\mu,v,j}$ where the two classes are prominently separated. After the projection, the dimension of histogram is reduced to one. To find the project direction is equivalent to find a hyperplane best separating the two classes because the project direction is perpendicular to the hyperplane. As we usually know, linear discriminant analysis (LDA), as a successful dimensionality reduction technique, can help to find the hyperplane and thus the project direction. Its objective is to find a projection A that maximizes the ratio of between-class scatter S_b against within-class scatter S_w [11]:

$$\arg \max_{A} \frac{|AS_b A^T|}{|AS_w A^T|}, \qquad (10)$$

where S_b and S_w are given by:

$$S_b = (M_{\mu,v,j}^{+} - M_{\mu,v,j})(M_{\mu,v,j}^{+} - M_{\mu,v,j})^T \\ + (M_{\mu,v,j}^{-} - M_{\mu,v,j})(M_{\mu,v,j}^{-} - M_{\mu,v,j})^T,$$

$$S_w = \sum_{H_{\mu,v,j}^{(k)} \in \mathcal{T}_{\mu,v,j}^{+}} (H_{\mu,v,j}^{(k)} - M_{\mu,v,j}^{+})(H_{\mu,v,j}^{(k)} - M_{\mu,v,j}^{+})^T \\ + \sum_{H_{\mu,v,j}^{(k)} \in \mathcal{T}_{\mu,v,j}^{-}} (H_{\mu,v,j}^{(k)} - M_{\mu,v,j}^{-})(H_{\mu,v,j}^{(k)} - M_{\mu,v,j}^{-})^T,$$

where

$$M_{\mu,v,j} = \frac{pM_{\mu,v,j}^{+} + nM_{\mu,v,j}^{-}}{p+n}$$

is the center of the whole $\mathcal{T}_{\mu,v,j}$.

The solution for the optimal problem stated in equation (10) is given as follows [16]:

$$A = S_w^{-1}(M_{\mu,v,j}^{+} - M_{\mu,v,j}^{-}). \qquad (11)$$

Actually, the magnitude of A is of no real significance because it merely scales the vectors projected into the subspace. Therefore, $D_{\mu,v,j}$ is given as the unit vector in the direction of A:

$$D_{\mu,v,j} = \frac{S_w^{-1}(M_{\mu,v,j}^{+} - M_{\mu,v,j}^{-})}{\|S_w^{-1}(M_{\mu,v,j}^{+} - M_{\mu,v,j}^{-})\|}. \qquad (12)$$

B. LGBMP-CCL

In Equation (12), we have to compute the within-class matrix S_w and its inverse. Furthermore, if S_w is singular, we have to project the histograms onto the directions of S_w's non-zero eigenvalues before applying LDA. This may consume much time when the training set is very large. We can simply project the histograms onto the center connecting line (CCL), where the project direction is represented as the unit vector pointing from one class center to the other, and expressed as:

$$D_{\mu,v,j} = \frac{M_{\mu,v,j}^{+} - M_{\mu,v,j}^{-}}{\|M_{\mu,v,j}^{+} - M_{\mu,v,j}^{-}\|}. \qquad (13)$$

Due to its compact form, LGBMP-CCL is more simple and easier to implement since it needs only to calculate the means of each class. Comparing equation (12) with (13), we can see that the only difference is a transformation matrix S_w^{-1}. This can be further seen clearly in Fig. 3: the solid points represent samples of one class while the hollow points of the other. The LGBMP-CCL and LGBMP-LDA solid lines denote the two project directions. Note that, under certain circumstances, the project direction of LGBMP-LDA and LGBMP-CCL will coincide.

Each 59-bin histogram $H_{\mu,v,j}^{(k)}$ belonging to $\mathcal{T}_{\mu,v,j}$ is mapped into a single value, namely $val_{\mu,v,j}^{(k)}$. The mapping formula is given by:

$$val_{\mu,v,j}^{(k)} = D_{\mu,v,j}^{T} \cdot H_{\mu,v,j}^{(k)}. \qquad (14)$$

Finally, all the single values are fitted to form the new feature vector:

$$V_{new}^{(k)} = \{val_{0,0,0}^{(k)}, \ldots, val_{0,0,m-1}^{(k)}, \\ val_{0,1,m-1}^{(k)}, \ldots, val_{7,4,m-1}^{(k)}\}. \qquad (15)$$

So far, we have mapped each 59-bin histogram to a single value. Therefore the total dimension of the feature vector has been reduced to $40m$, which is less than that of LBP and LGBP feature.
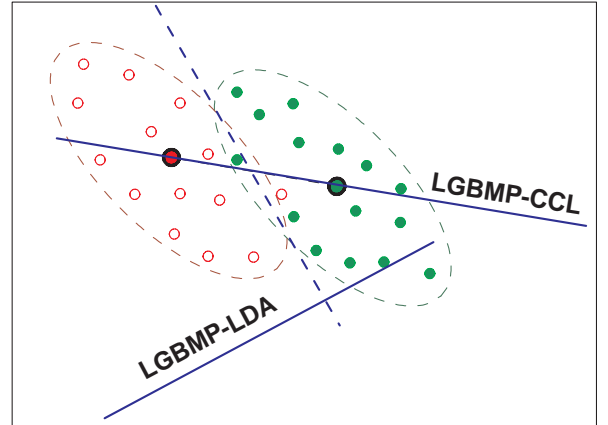

Fig. 3. The project directions of LGBMP-LDA and LGBMP-CCL.

IV. EXPERIMENTS

In order to illustrate the effectiveness of the proposed approaches for multi-view gender classification, we have conducted comparative experiments on the CAS-PEAL face database [13]. This is a large-scale face database that currently contains 30,864 facial images of 1,040 individuals with 9 different poses including looking up pose, looking middle pose and looking down pose with 3 different azimuths, namely 0 degree, 15 degree and 30 degree, respectively.

Before the facial images are available for feature extraction, they have to be pre-processed. The eye coordinates of each image are first manually located. Then the image is cropped by the eye coordinates and normalized to the size of 130×150 pixels. After that, the histogram is equalized. For each pose we pick 200 images of male and female respectively for training and the rest are used for test. Table I shows the detailed contents of our data sets where the meaning of each data set's name 'PX_YY' is illustrated as follows: 'P' is the acronym for 'Pose', 'X' represents the subject's pose, facing down (D), middle (M) or up (U), and 'YY' indicates the azimuth from which the picture is taken, either 0, 15 or 30 degree. The total number of training sample is 3,600 and the total number of test sample is 10,784.

TABLE I

THE CAS-PEAL FACE DATABASE FOR MULTI-VIEW GENDER CLASSIFICATION. THE NAME OF DATA SET CONSISTS OF TWO PARTS: 'PD','PM' AND 'PU' DENOTE THREE DIFFERENT TYPES OF FACE POSE—FACING DOWN, MIDDLE AND UP; THE NUMBER DENOTES THE POSE AZIMUTH—$0°$, $15°$ AND $30°$

Data Set	No. Male	No. Female	No. Total data	No. Training	No. Test
PD_00	445	595	1,040	200×2	640
PD_15	846	1,032	1,878	200×2	1,478
PD_30	846	1,032	1,878	200×2	1,478
PM_00	445	595	1,040	200×2	640
PM_15	844	1,032	1,876	200×2	1,476
PM_30	844	1,032	1,876	200×2	1,476
PU_00	445	595	1,040	200×2	640
PU_15	846	1,032	1,878	200×2	1,478
PU_30	846	1,032	1,878	200×2	1,478
Total	6,407	7,977	14,384	3,600	10,784

In all our experiments, the image is first convolved by $5 \times 8 = 40$ Gabor filters and $LBP_{8,1}^{u2}$ operator is used to encode the GMPs. After that, the LGBP images are divided into $K \times K$ equal sized rectangular regions where K is ranged from 5 to 10 (some experiments use K ranged from 7 to 10). Then the 59-bin histogram of each region are mapped into a single value and these values together compose the final feature vector. All the SVMs in our experiments come from LibSVM 2.84 [12] and the model parameters are obtained by cross-validation.

A. LGBMP-LDA versus LGBMP-CCL

As stated in section III, two solutions have been proposed for dimensionality reduction, the LGBMP-LDA and LGBMP-CCL. We set up a comparative experiment between the two solutions where K is ranged from 5 to 10 and three different SVM kernels, RBF, linear and three-degree polynomial kernel, are used. We evaluate the two solutions on all test sets and compute a total average accuracy. The results are shown in Table II. From this table, it can be seen that LGBMP-CCL achieves a little higher average accuracy than LGBMP-LDA in most cases. The largest accuracy difference is about 1.5% when $K = 9$ and RBF kernel is used. In principle, dimensionality reduction using LDA is more likely to achieve the local optimal results however it does not inevitably lead to the best final accuracy, which also depends on the classifier. More particularly, LGBMP-LDA might overfits to the data distribution due to some noise. As LGBMP-CCL is more simple, easier to implement and with lower complexity, we only use LGBMP-CCL in the subsequent experiments to compare with other methods: SVMs+Grayscale pixel, SVMs+Gabor and SVMs+LBP method. We can see that although LGBMP-LDA has a little lower accuracy than LGBMP-CCL, it still outperforms the other methods and its feature dimension is exactly the same as LGBMP-CCL. Additionally, total training time for the SVMs with all the six values of K is more than one day on a Lenovo cluster which consists of three fat nodes and thirty thin nodes. We only use a single fat node which has 32G RAM and two 3.2GHz Intel(R) Xeon(TM) four-core CPUs.

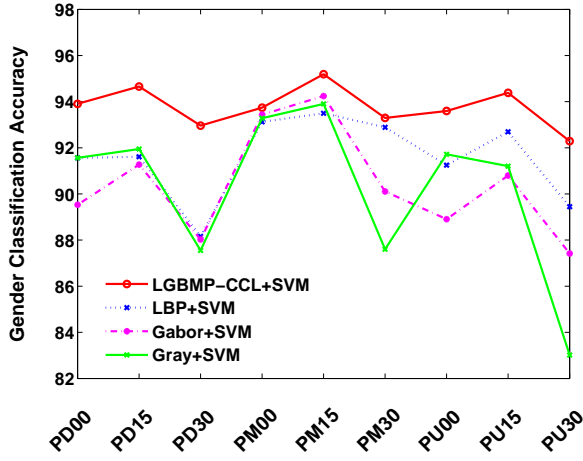### B. Comparisons based on different face poses

In this part, we will compare the proposed feature extraction method LGBMP-CCL with support gray face method [1], support LBP face method [4] and support Gabor face method [5] based on all 9 face poses. The most important difference between our work and the existing approaches is that we combine Gabor wavelet transformation, LBP and histogram mapping while they use only one of them or even use merely the gray-scale pixel.

We only use the RBF kernel in this part to compare the individual accuracy on test set of each face pose respectively. The experimental results are shown in Fig. 4. The four subfigures are drawn with respect to K which is ranged from 7 to 10 in this part. From these figures, it is obvious that the classification accuracy of LGBMP-CCL is higher than that of gray-scale pixel, Gabor and LBP method on all test sets. A highest rate of 96.71% has been achieved on the PM_00 test set when $K = 10$. The gray-scale method almost fails with a lowest accuracy of 83.02% on the PU_30 data set while LGBMP-CCL greatly increases the accuracy to 93.16%. Comparing to the LBP method, LGBMP-CCL also makes an accuracy improvement of $1\% \sim 2\%$. If we fix a face pose and evaluate the accuracy on different azimuth, we can see that LGBMP-CCL is more stable to the azimuth variation than the other methods. Similar conclusion can be drawn by fixing the azimuth while varying the face pose. We can also see how the block number $K \times K$ influences the final performance. Take the test results on PM_00 for instance, the accuracy keeps increasing from 93.74% to 96.71% in accordance with K. The accuracy on 30 degree
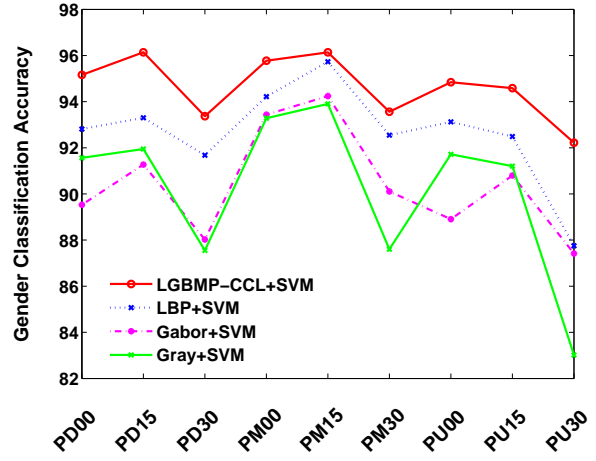
TABLE II

AVERAGE ACCURACY (%) EVALUATED USING LGBMP-LDA AND LGBMP-CCL ON CAS-PEAL DATABASE. EACH GABOR FEATURE MAP IS DIVIDED INTO $K \times K$ BLOCKS. K IS RANGED FROM 5 TO 10.
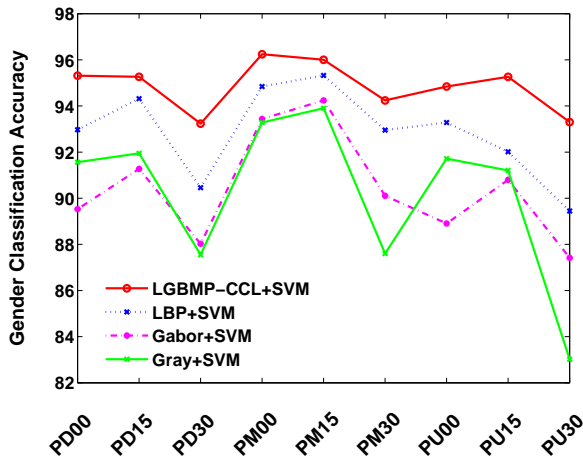
Kernel	Method	5×5	6×6	7×7	8×8	9×9	10×10
RBF	LGBMP-CCL	**92.99**	**93.20**	**93.79**	**94.50**	**94.71**	**94.96**
	LGBMP-LDA	92.95	93.05	93.33	93.68	93.15	93.59
Linear	LGBMP-CCL	91.18	**92.21**	92.17	**93.54**	**93.66**	**94.09**
	LGBMP-LDA	**91.68**	91.93	**92.63**	93.24	92.35	93.23
Poly	LGBMP-CCL	92.46	92.86	**93.29**	**93.92**	**94.08**	**94.61**
	LGBMP-LDA	**92.83**	**92.95**	93.11	93.64	93.08	93.58

(a)

(b)

(c)

(d)

Fig. 4. The comparative results of LGBMP, LBP, Gabor and gray-scale methods with RBF kernel SVMs and $K \times K$ blocks for each face image. (a) $K = 7$. (b) $K = 8$. (c) $K = 9$. (d) $K = 10$.

data set also has an improvement of 1% from $K = 7$ to $K = 10$. This demonstrates that the spatial information of facial image is very important to multi-view gender classification problem. Note that, we can of course keep on increasing the region number and dividing the LGBP images into more regions nevertheless the accuracy will advance very little while the dimension increase greatly. The reason why LGBMP outperforms the others is illustrated as follows: Gabor filters is robust to variations of illumination and pose, therefore they can remove a lot of noise. Meanwhile, LBP has been proved capable of drawing local shape and texture information effectively. Furthermore, the spatial information can be appropriately obtained by using spatial histogram. So, if we combine these functions, we naturally receive a good result.
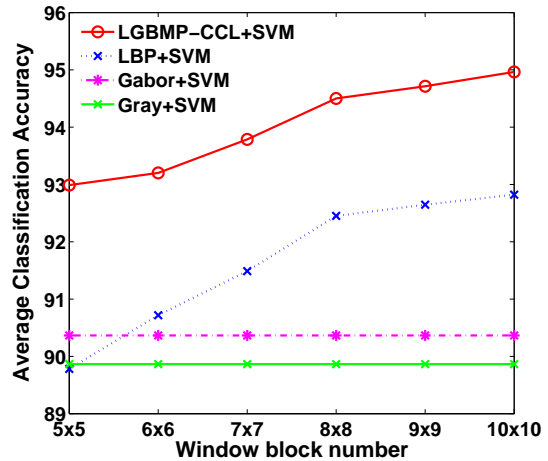
C. Comparisons based on different kernels

To further validate the effectiveness of LGBMP-CCL, we compare its performance with other mentioned methods using three different SVM kernels (RBF, linear and three-degree polynomial kernel) and different K values ranged from 5 to 10 on all 9 face poses. We do not evaluate the accuracy on each individual test set, instead, we make a total statistics. Fig. 5 shows the experimental results. From these figures, we can see that LGBMP-CCL occupies the highest correct rates among all the approaches using all the three SVM kernels. Therefore, LGBMP-CCL is stable and robust to different SVM kernels and proved to be a good face representation approach. The reason why the linear kernel performs as well as the nonlinear ones might be that the data set is almost linear-separable in the feature space.
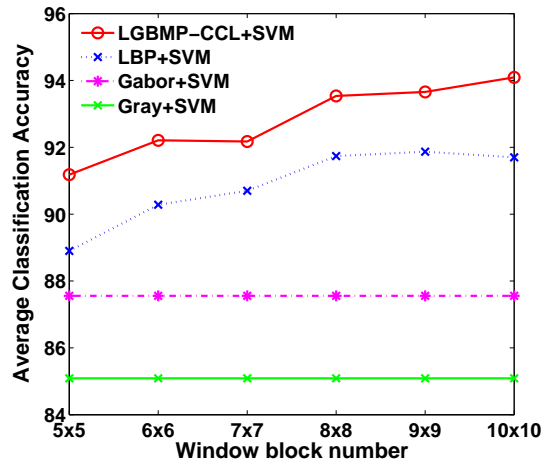
D. Feature dimension

We also compare the dimension of feature vector among different feature extraction methods. Fig. 6 shows the feature dimension generated by LGBMP methods (LGBMP-LDA and LGBMP-CCL have the same feature dimension, therefore in this part we notate them together as LGBMP for short), LBP method and gray-scale pixel method. From this figure, we can see that the dimension of LGBMP methods is the least among the three. Gabor method has a dimension of 468,000 [4] which is much larger than LGBMP methods and thus we do not plot that line.
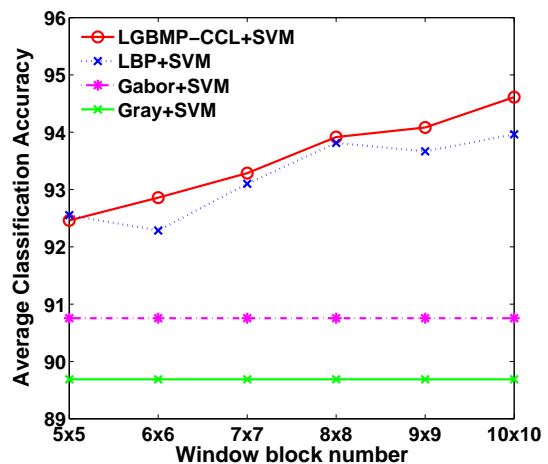
V. CONCLUSION AND DISCUSSION

In this paper, we proposed a new feature extraction approach LGBMP. This approach is impressively insensitive to appearance variations due to illumination and face pose. Meanwhile, two histogram mapping solutions, LGBMP-LDA and LGBMP-CCL are employed in LGBMP. The effectiveness of LGBMP comes from several aspects including the multi-scale and multi-orientation Gabor wavelets decomposition, the uniform local binary pattern, and the mapping function for each local spatial histogram. Therefore, LGBMP consists of many information corresponding to different face components at different scales and orientations.

Fig. 5. Average classification accuracy (%) comparison with three different SVM kernels. (a) RBF kernel. (b) Linear kernel. (c) Polynomial kernel.

Experimental evaluations of our proposed approaches have evidently illustrated the effectiveness and robustness of LGBMP to the general variations of illumination and face pose. On the CAS-PEAL face database, our proposed method has achieved the best performance with the lowest feature dimension over all the other methods.

Future efforts will be focused on how to add phase information of multi-scale and multi-orientation Gabor wavelets into our proposed method. At the same time, to further improve the classification accuracy, we try to find other histogram mapping functions other than LGBMP-CCL and LGBMP-LDA. Furthermore, we try to apply LGBMP feature to other computer vision domain, such as face recognition.
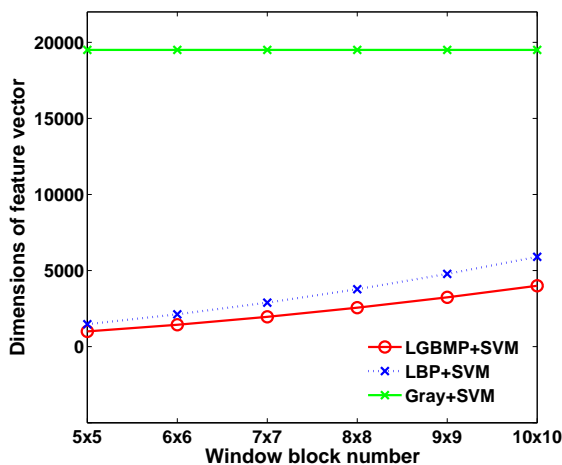
Fig. 6. Comparison of feature dimension.

REFERENCES

[1] B. Moghaddam and M. H. Yang, "Learning gender with support faces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 707-711, 2002.

[2] K. Balci and V. Atalay, "PCA for gender estimation: which eigenvectors contribute," *Proceedings of the Sixteenth International Conference on Pattern Recognition*, vol. 3, pp. 363-366, 2002.

[3] S. Hosoi, E. Takikawa and M. Kawade, "Ethnicity estimation with facial images," *Proceedings of the Sixth IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 195-200, 2004.

[4] H.C. Lian and B.L. Lu, "Multi-view gender classification using local binary patterns and support vector machines," *Proceedings of the Third International Symposium on Neural Networks*, pp. 202-209, 2006.

[5] R. Iga, K. Izumi, H. Hayashi, G. Fukano and T. Ohtani, "A gender and age estimation system from face images," *SICE 2003 Annual Conference*, vol. 1, 2003.

[6] M.Lades, J.C. Vorbruggen, J. Buhmann, J. Lange, C. Malsburg, R.P. Wurtz, W. Konen, "Distortion invariant object recognition in the dynamic link architecture," *IEEE Transactions on Computers*, pp. 300-311, 1993.

[7] T. Ojala, M. Pietikainen and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, 2002.

[8] T. Ahonen, A. Hadid and M. Pietikainen, "Face recognition with local binary patterns," *ECCV 2004*, pp. 469-481, 2004.

[9] 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," *Proceedings of the Tenth IEEE International Conference on Computer Vision*, pp. 786-791, 2005.

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

[11] A. Martinez and A. Kak, "PCA versus LDA," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, 2001.

[12] C.C. Chang and C.J. Lin, "LIBSVM: a library for support vector machines," *Software available at http://www.csie.ntu.edu.tw/cjlin/libsvm*, vol. 80, pp. 604-611, 2001.

[13] W. Gao, B. Cao, S. Shan, D. Zhou, X. Zhang and D. Zhao, "The CAS-PEAL large-scale chinese face database and baseline evaluations," *Technical report of JDL, http://www.jdl.ac.cn/peal/peal tr.pdf*, 2004.

[14] G.W. Cottrell, "Empath: face, emotion, and gender recognition using holons," *Advances in Neural Information Processing Systems 3*, R.P. Lippmann, J.E. Moody, and D.S. Touretzky, eds. pp. 564-571, 1991.

[15] R. Brunelli and T. Poggio, "Hyperbf networks for gender classification," *Proc. DARPA Image Understanding Workshop*, pp. 311-314, 1992.

[16] R.O. Duda, P.E. Hart and D.G. Stork, *Pattern Classification*. Wiley-Interscience, 2000.

[17] H.C. Lian and B.L. Lu, "Multi-view gender classification using multi-resolution local binary patterns and support vector machines," *accepted by Internationl Jounal of Neural Systems*.