Neurocomputing 165 (2015) 190–201

Contents lists available at ScienceDirect

Neurocomputing

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

Patterns of Weber magnitude and orientation for uncontrolled face representation and recognition ☆

Yinyan Jiang [a,b], Biao Wang [a,b], Yicong Zhou [c], Weifeng Li [a,b,*], Qingmin Liao [a,b]

[a] Department of Electronic Engineering/Graduate School at Shenzhen, Tsinghua University, China
[b] Shenzhen Key Laboratory of Information Science and Technology, Guangdong, China
[c] Department of Computer and Information Science, University of Macau, Macau, China

ARTICLE INFO

Article history:
Received 14 November 2014
Received in revised form
19 February 2015
Accepted 2 March 2015
Communicated by Ran He
Available online 25 March 2015

Keywords:
Face recognition
Weber's law
Feature extraction
Local binary patterns (LBP)
Local XOR patterns (LXP)
Fisher's linear discriminant (FLD)

ABSTRACT

Robust, discriminative and computationally efficient feature extraction is vital for a successful real-world face recognition system. To this end, we present a novel local feature descriptor, named patterns of Weber magnitude and orientation (PWMO), for face representation and recognition. Instead of merely taking advantage of the pixel intensity which is sensitive to variant impact factors such as illumination variations and noises, we describe a pixel with two robust statistic attributes of the local patch centered around it: the histogram of Weber magnitude and the dominant Weber orientation. By encoding them in a self-similarity manner with patch-based local binary patterns (p-LBP) and patch-based local XOR patterns (p-LXP) respectively, we obtain a robust representation for face images. To further enhance the discriminative power, we extend PWMO to its multi-scale version, and apply the block-based Fisher's linear discriminant (BFLD) to reduce the dimensionality and select the most discriminative features. The Fisher separation criterion (FSC) based block weighting scheme is incorporated for discriminative classification. We evaluate the proposed face representation method on two publicly available face databases: FERET and FRGC version 2.0 experiment 4 (FRGC-204). The recognition results demonstrate that the proposed method performs much better than most of the state-of-the-arts, and achieves comparable recognition performance with the recently proposed state-of-the-art algorithm based on the fusion of Gabor magnitude and phase, with only 1/7 storage requirement and 1/10 computational cost.

1. Introduction

Due to the broad prospect in real-world applications such as surveillance, biometrics, human computer interaction and image retrieval, automatic face recognition remains an extensively studied topic in computer vision communities. During the last several decades, hundreds of effective technologies have been proposed to address this problem [1,2], and the performance evaluation results on large-scale face databases have demonstrated that machine vision can surpass human vision under well-controlled circumstances [3]. However, face recognition under uncontrolled circumstances is still challenging and remains an open problem, due to large intrapersonal variations which largely impact the appearances of face images, such as expression, illumination, pose, noise, occlusion, and aging.

As in any pattern classification tasks, efficient feature extraction is acknowledged to be vital for designing a successful face recognition system. Basically, existing feature extraction technologies for face representation could be divided into two categories: holistic feature extraction and local feature extraction. Holistic feature extraction usually bases on the subspace learning or spatial-frequency transformation technologies. The former exploits the whole face images to construct either a linear subspace by subspace learning methods such as principal component analysis (PCA) [4], Fisher's linear discriminant (FLD) [5], independent component analysis (ICA) [6], local preserving projection (LPP) [7], unsupervised discriminant projection (UDP) [8], or a nonlinear subspace by their kernel counterparts [9]. The latter utilizes the spatial-frequency transformation technologies such as Fourier transform [10], wavelets [11], and discrete cosine transform [12].

Among variant local feature descriptors, local binary patterns (LBP) and Gabor wavelets have been recognized as the most successful ones [13]. LBP is a computationally efficient nonparametric descriptor originally designed for texture analysis [14], which encodes the local patterns in a self-similarity manner by comparing the central pixel

☆This work was partially supported by the Macau Science and Technology Development Fund under Grant 017/2012/A1 and by the Research Committee at University of Macau under Grants SRG007-FST12-ZYC, MYRG113(Y1-L3)-FST12-ZYC and MRG001/ZYC/2013/FST.

* Corresponding author at: Department of Electronic Engineering/Graduate School at Shenzhen, Tsinghua University, China.
E-mail addresses: jiangyy12@mails.tsinghua.edu.cn (Y. Jiang), wangbiao08@mails.thu.edu.cn (B. Wang), yicongzhou@umac.mo (Y. Zhou), Li.Weifeng@sz.tsinghua.edu.cn (W. Li), liaoqm@tsinghua.edu.cn (Q. Liao).

http://dx.doi.org/10.1016/j.neucom.2015.03.009
0925-2312/

with its neighbors and can well describe the local image structures. Recently, inspired by the pioneering work and its remarkable results on FERET face database provided by Ahonen et al. [15], face representation based on the spatial histogram model of LBP, which can well capture both the structure and texture information, arouses increasing interest and becomes one of the most popular and successful technologies in face recognition applications. In the last few years, numeric LBP variants have been proposed and successfully applied to face recognition tasks [16]. For example, Tan and Triggs [17] proposed local ternary patterns (LTP) to alleviate the sensitivity to the noises and near-uniform image regions. Local derivative pattern (LDP) [18] is a generalization of LBP which aims at capturing higher-order local information. Three-Patch LBP (TPLBP) and Four-Patch LBP (FPLBP) were proposed by Wolf et al. [19] to explore and encode similarities between neighboring image patches rather than neighboring pixels. Recently, semantic pixel sets LBP (spsLBP) [20] obtains histograms of semantic pixel sets based LBP with a robust code voting. However, most LBP variants are inadequate for nonmonotonic illumination changes.

The Gabor wavelets resemble the receptive fields of simple cells in mammalian's visual cortex, which are spatially localized and selective to spatial orientations, and they have been theoretically proved to be optimally localized in the space and frequency domains [21]. The Gabor feature and its variants have been extensively and successfully used in face recognition [22,23]. Compared with LBP, Gabor feature is more robust to several variations such as illumination and noise. However, the application of real-time face recognition is restricted due to the time-consuming feature extraction process by convolving a face image with Gabor filter banks and the relatively high feature dimensionality. For example, Gabor feature extraction with filter banks of 5 scales and 8 orientations as typically used would require 40 convolutions. As for a 128×128 image, the dimensionality of its Gabor feature vector would be up to $40 \times 128 \times 128 = 655{,}360$.

Recently, some researchers attempted to apply LBP on stable pixel attributes rather than the pixel intensity to seek more effective face representations. For example, Zhang et al. [24] presented local Gabor binary patterns (LGBP) by applying LBP on Gabor magnitudes. Zhang et al. [25] proposed local Gabor XOR patterns (LGXP) by performing on Gabor phases local XOR patterns (LXP), which is a variant of LBP. Xie et al. [26] introduced the fusion of LGBP and LGXP and demonstrated its impressive results on FERET and FRGC-204. However, these methods still suffer from the disadvantage of high computational cost, which motivates researchers to seek more efficient local descriptors for face representation. Very recently, Vu and Caplier [27] attempted to extend the pixel-based self-similarity in LBP to patch-based self-similarity, in which a central pixel is represented by the accumulated gradient magnitudes across different directions with the help of the histogram of gradient (HOG) in a local patch centered around it. The proposed descriptor, named patterns of oriented edge magnitude (POEM), achieved impressive performance on FERET and LFW face databases, and was declared to be the first computationally efficient descriptor with comparable performance to its Gabor's counterparts. However, as indicated in [28], the gradient magnitude is sensitive to illumination variations. Our further experiments on FRGC-204 in Section 3 demonstrate the inferior performance of POEM compared with those based on Gabor features in poor illumination and uncontrolled circumstances.

To develop a successful face recognition system for real-time applications, it is important to evaluate which feature extractors are suitable. Generally speaking, a well-designed feature extractor should meet the following three criteria [27]: (1) *robustness*, which means it should minish the intraperson variations; (2) *distinctiveness*, which means it should retain enough discriminative power for interperson variations; and (3) *high computational efficiency*, which requires the feature extraction process should not be time-consuming and the final face representation features should not be storage-consuming. Compared with holistic feature extraction, it is reported by Heisele et al. [29] that local feature extraction improves the recognition rates by 60%, probably due to the robustness to variations of facial expression, illumination and occlusion, etc. As for local feature extractors, to the best of our knowledge, on large-scale challenging face databases such as FRGC-204, no other existing descriptors can obtain better or comparable performance than those based on Gabor features with much less computational cost.

Based on a preliminary conference version of this paper [30], we propose a local descriptor which satisfies the three aforementioned criteria, named patterns of Weber magnitude and orientation (PWMO), for effective face representation with efficient computations for face recognition. In this work, we extensively analyze and illustrate the effectiveness and robustness of PWMO as face representation, and extend it to a multiscale version for further improvement. We further adopt a block-based Fisher linear discriminant (BFLD) for discriminative feature selection, which exploits Fisher separation criterion as block weighting scheme. The superiority of PWMO comes from several aspects: (1) instead of the pixel intensity, two human perception inspired illumination- and noise-insensitive statistics of a local patch, i.e., the histogram of Weber magnitude (HWM) and the dominant Weber orientation (DWO), are applied to represent the central pixel, thus making PWMO robust to illumination and noise variations; (2) the local information for face images are encoded in a patch-based self-similarity manner by performing patch-based LBP (p-LBP) and patch-based LXP (p-LXP) on HWM and DWO, respectively, thus making PWMO robust to pose variation within certain degrees; (3) the spatial histogram model is utilized for face representation, thus retaining the structure information and making PWMO robust to small registration errors; (4) multi-scale encoding scheme is utilized to capture both local and more global structures; and (5) both the calculation of HWM and DWO and the encoding processes are computationally efficient. Furthermore, we take advantage of block-based Fisher's linear discriminant (BFLD) to select the most discriminative compact features, and the block weighting scheme based on Fisher separation criterion (FSC) is proposed for discriminant classification. Extensive experiments are conducted on three large face databases for evaluating the performance and efficiency of the proposed method, and the results demonstrate that it is comparable to the recently proposed algorithm based on the fusion of Gabor magnitude and phase, which has much higher computational efficiency.

The remainder of this paper is organized as follows. Section 2 briefly describes the definition of Weber magnitude and orientation, and proposes a generalized formulation of p-LBP and p-LXP, and then details the proposed PWMO feature extraction process. In Section 3, experimental results and analysis are provided to illustrate the effectiveness and efficiency of our method. Section 4 concludes the paper.

2. Patterns of Weber magnitude and orientation (PWMO) for face representation

2.1. Weber magnitude and orientation

Weber's law suggests that for a stimulus, the ratio between the smallest perceptual change and the background is a constant, which implies that stimuli are perceived not in absolute terms but in relative terms. Inspired from this, Chen et al. [31] proposed a local descriptor named Weber local descriptor (WLD) for the task of texture classification, which consists of two components: differential excitation and orientation. The former describes the relative intensity differences of a central pixel against its neighbors, and the latter is the gradient orientation of the central pixel. The two could provide complementary information for local

pattern description, and WLD has empirically proved to be more effective than LBP, SIFT and Gabor for texture classification tasks [31].

In our previous work, we theoretically proved the differential excitation part is illumination insensitive and proposed an efficient illumination-insensitive representation approach (namely, Weberface [32]) for face images under uneven illumination conditions. The orientation part is also illumination insensitive according to the Gradientface approach proposed by Zhang et al. [33]. Both Weberface and Gradientface take advantage of Gaussian filtering as a preprocessing step to alleviate the side-effect of noise and shadow, and this step proves to be important for robust face recognition.

We define Weber magnitude and orientation in this subsection. As mentioned above, for a face image I, Gaussian filtering is first performed as a preprocessing step:

$$I' = I * G(x, y, \sigma), \tag{1}$$

where $*$ is the convolution operator and

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \tag{2}$$

is the Gaussian kernel function with standard deviation σ.

Weber magnitude is defined as follows:

$$\textit{Weber magnitude} : \xi_m(x_c) = \arctan\left(\alpha \sum_{i=0}^{p-1} \frac{x_c - x_i}{x_c}\right), \tag{3}$$

where the arctangent function is used to prevent the range of the output from being too large and thus could partially suppress the side-effect of noise. x_c denotes the center pixel, and the set $\{x_i, i = 0, 1, \ldots, p-1\}$ are the neighboring pixels as illustrated in Fig. 1(a). p is the number of neighbors, and α is a parameter for adjusting (magnifying or shrinking) the intensity difference between neighboring pixels. Note that $\xi_m(x_c)$ ranges in $[-\pi/2, \pi/2]$. If $\xi_m(x_c)$ is positive, it implies that central pixel is lighter than the surroundings. If $\xi_m(x_c)$ is negative, the central pixel is darker than the surroundings. If $\xi_m(x_c)$ is close to zero, then it mainly corresponds to a flat area.

As shown in Fig. 2, we plot an average histogram of Weber magnitude over 10,000 face images. One can find that there are more frequencies at the center of the average histogram (i.e., $[-\pi/6, \pi/6]$). It is mainly due to the relatively large area of flat regions in face images such as forehead and cheeks. Although there are less frequencies at the two sides of the average histogram (i.e., $[-\pi/2, -\pi/3]$ and $[\pi/3, \pi/2]$), they correspond to salient regions such as eyebrows, eyes, lips, and noses, which are important for face recognition.

Weber orientation is defined as follows:

$$\textit{Weber orientation} : \xi_o(x_c) = \arctan\left(\frac{x_1 - x_5}{x_3 - x_7}\right), \tag{4}$$

where $x_1 - x_5$ and $x_3 - x_7$ indicate the intensity difference in x and y direction, respectively. We rescale $\xi_o(x_c)$ to the range of $[0, 2\pi]$. Fig. 3 shows the average histogram of Weber orientation over 10,000 face images. One can see that the frequencies are relatively uniform except two peaks at around $\pi/2$ and $3\pi/2$ due to the delimitation effect of the arctangent function.

Fig. 1(b)–(d) illustrates two sample face images, their corresponding Weber magnitudes and Weber orientations. As can be seen, both Weber magnitude and orientation are insensitive to illumination variations. Another significant merit of Weber magnitude and orientation is their robustness to noises. It results from two factors: one is directly from the noise removal effect of the preprocessing by Gaussian filter (Eq. (1)); the other can be easily concluded from the definitions of Weber magnitude (Eq. (3)) and Weber orientation (Eq. (4)), including two aspects: (a) because they are both defined in the form of ratio images, Weber magnitude and orientation are robust to multiplicative noise; (b) the gradient calculation and the arctangent function tend to suppress the side-effect of noises.

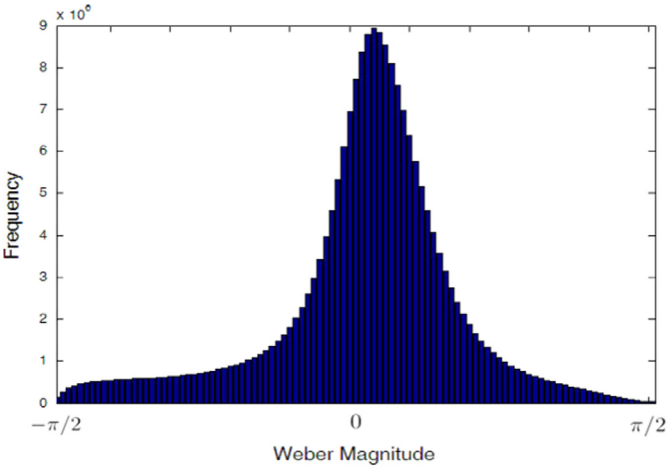


Fig. 2. The average distribution of Weber magnitude over 10,000 face images.

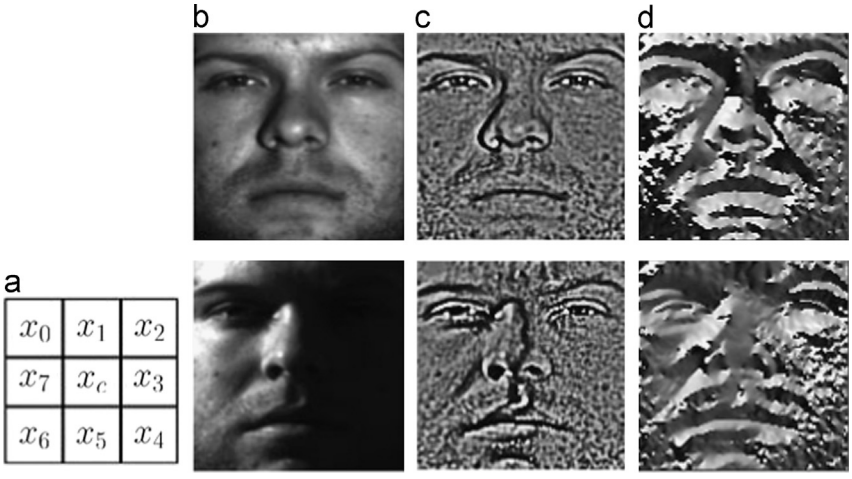


Fig. 1. (a) Illustration of a pixel and its eight neighbors; (b) original images; (c) Weber magnitudes; (d) Weber orientations.

2.2. Generalized patch-based LBP (p-LBP) and LXP (p-LXP)

Face representation by the spatial histogram of LBP can encode both texture and structure information, and thus obtains impressive results by performing on gray image [15], Gabor magnitude [24,26] and Gabor phase [25,26]. Traditional LBP encodes each pixel in a self-similarity manner with eight bit code by comparing the pixel intensity with that of its neighbors. Formally, it could be described as follows:

$$LBP(x_c,y_c)=\sum_{n=0}^{7}2^n s(I_c-I_n), \tag{5}$$

in which (x_c,y_c) is the location of the central pixel. I_c and I_n are the intensity of the central pixel and its n-th neighbor, and $s(u)$ is 1 for $u\geq 0$ and 0 otherwise. However, the pixel intensity is sensitive to nonmonotonic illumination and pose variations. In this subsection, we further extend it to the generalized patch-based LBP (p-LBP), which encodes each pixel by comparing certain statistic attributes of the neighboring patch with that of the patch centered at the current pixel. Basically, p-LBP can be formulated as follows:

$$\text{p}-\text{LBP}(x_c,y_c)=\left[\sum_{n=0}^{7}2^n D(A_1(P_c^r),A_1(P_{c_n^R}^r)),\ldots,\sum_{n=0}^{7}2^n D(A_q(P_c^r),A_q(P_{c_n^R}^r))\right], \tag{6}$$

in which (x_c,y_c) is the location of the central pixel. P_c^r is the image patch centered at (x_c,y_c) with radius equal to r, and c_n^R $(n=0,1,\ldots,7)$ are the neighboring pixels which are uniformly distributed on a ring of radius R around (x_c,y_c). $A(P_c^r)=[A_1(P_c^r)\ A_2(P_c^r)\ \cdots\ A_q(P_c^r)]$ is the predefined statistic attributes for image patch P_c^r, and q is the number of attributes. Here note that $A(P_c^r)$ could be either a scalar (e.g., average intensity as in [34]) or a vector (e.g., histogram of gradient orientation as in [27]). The similarity measure is defined by

$$D(A_i(P_c^r),A_i(P_{c_n^R}^r))=\begin{cases}1, & A_i(P_c^r)-A_i(P_{c_n^R}^r)\geq 0\\ 0 & \text{otherwise.}\end{cases} \tag{7}$$

The encoding process of p-LBP is illustrated in Fig. 4. We can easily conclude that traditional LBP descriptor is a special case of p-LBP by setting $r=0,R=1,A(P_c^r)=I_c$, while POEM [27] is also a special case of p-LBP by setting $r=3,R=5,A(P_c^r)=HOG(P_c^r)$, in which $HOG(P_c^r)$ is the histogram of gradients for image patch P_c^r. LXP was first proposed in [25] for encoding the Gabor phase in a self-similarity manner by comparing the quantized phase orientation of the central pixel with that of its neighboring pixels. We also extend it to p-LXP and formulate it in the same way as Eq. (6), except that the similarity measure is defined by

$$D(A_i(P_c^r),A_i(P_{c_n^R}^r))=\begin{cases}1, & A_i(P_c^r)\neq A_i(P_{c_n^R}^r)\\ 0 & \text{otherwise.}\end{cases} \tag{8}$$

The encoding process of p-LXP is illustrated in Fig. 5, and traditional LXP descriptor is a special case of p-LXP by setting $r=0,R=1,A(P_c^r)=GP(x_c,y_c)$, in which $GP(x_c,y_c)$ is the Gabor phase for the central pixel. Compared with traditional LBP, the defined p-LBP/LXP has the following merits if proper statistic attributes are adopted:

1. Instead of pixel intensity which is sensitive to illumination and noise variations, robust statistic attributes in a local patch are utilized to represent the central pixel, thus making p-LBP/LXP much more robust.
2. The patch-based self-similarity calculation can provide richer information about face images, and also make p-LBP/LXP less sensitive to pose variations.

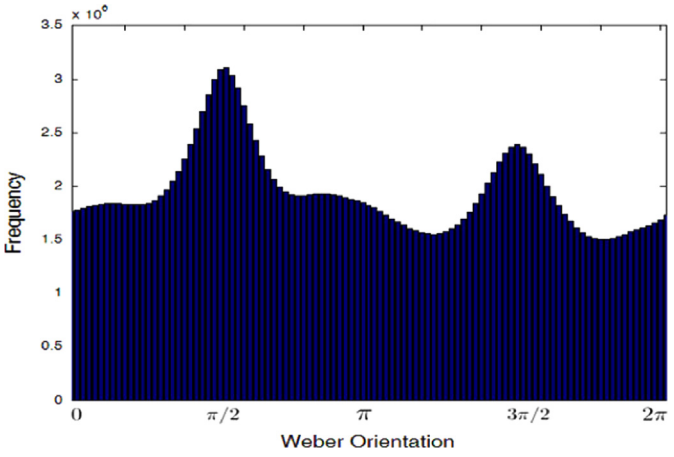


Fig. 3. The average distribution of Weber orientation over 10,000 face images.

2.3. PWMO feature extraction

In order to design a robust, discriminative and computationally efficient descriptor for face representation, we propose a novel descriptor following the framework of the generalized p-LBP/LXP. The proposed PWMO descriptor consists of two parts: patterns of Weber magnitude (PWM) and patterns of Weber orientation (PWO). For a specific face image, PWM encodes the histogram of Weber magnitude by p-LBP with $A(P_c^r)$ defined as follows:

$$A(P_c^r)=histogram\_wm(P_c^r), \tag{9}$$

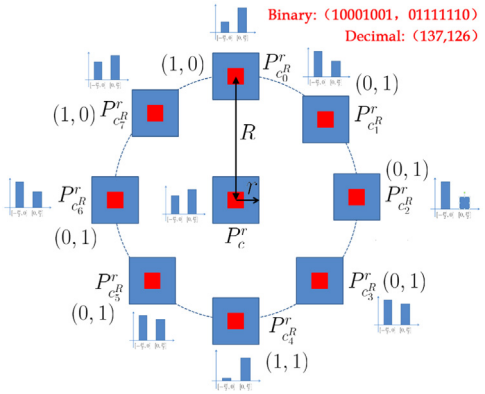


Fig. 4. Illustration of encoding process of p-LBP.

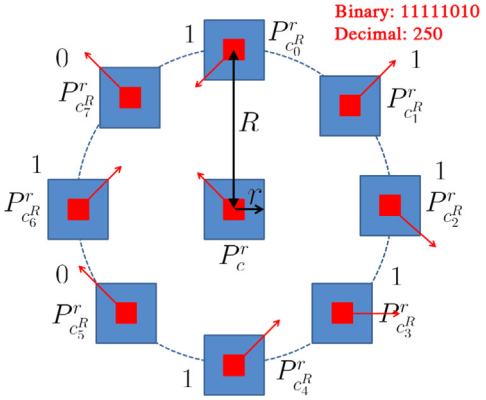


Fig. 5. Illustration of encoding process of p-LXP.

where $histogram\_wm(P_c^r)$ is the histogram of Weber magnitude for the patch centered at (x_c, y_c) with the radius equal to r. It is calculated by dividing the Weber magnitude ranging from $[-\pi/2, \pi/2]$ into bin_{wm} bins, and the vote weight for each bin is the summation of the absolute Weber magnitude.

PWO encodes the dominant Weber orientation by p-LXP with $A(P_c^r)$ defined as follows:

$$A(P_c^r) = dominant\_wo(P_c^r), \tag{10}$$

where $dominant\_wo(P_c^r)$ is the dominant Weber orientation with the largest accumulative magnitude across certain gradient orientation in the patch centered at (x_c, y_c) with the radius equal to r. The dominant orientation could be easily found with the help of the histogram of gradient orientation (calculated by Eq. (4)) and the gradient magnitude defined by

$$\xi_{om}(x_c) = \sqrt{\left(\frac{x_1 - x_5}{x_c}\right)^2 + \left(\frac{x_3 - x_7}{x_c}\right)^2}, \tag{11}$$

where the pixels x_c, x_1, x_3, x_5, x_7 are illustrated in Fig. 1(a), and the normalization by x_c is to suppress the impact of uneven illumination. We denote the number of orientation bins as bin_{wo}. An intuitive illustration of the derivation of $dominant\_wo(P_c^r)$ is illustrated in Fig. 6, where $bin_{wo} = 4$.

To represent the face image compactly by the spatial histogram of PWM and PWO, the uniform binary patterns are used. A LBP code is 'uniform' if it contains no more than two 0–1/1–0 transitions. For more details, refer to [15]. To encode both texture and structure information for human face, the code map of a face image is divided into several nonoverlapping blocks and histogram computed in each block is concatenated together to form the final representation. In our scheme, the spatial histograms of PWM and PWO are concatenated into a single feature vector, which is the proposed patterns of Weber magnitude and orientation (PWMO). The framework of face representation based on PWMO is illustrated in Fig. 7, in which $bin_{wm} = 2$, $bin_{wo} = 4$. Compared with POEM [27], which utilizes HOG of the local patch as the statistic attribute of the central pixel, the proposed PWMO takes advantage of the histogram of Weber magnitude and the dominant Weber orientation. Both of them prove to be much more robust [32,33] and discriminative for local pattern description than SIFT, LBP and Gabor [31], thus expecting better representation ability.

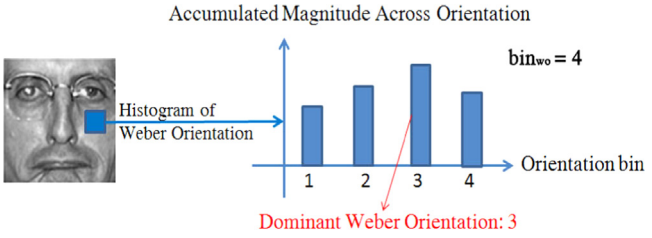


Fig. 6. Illustration of the calculation of the dominant Weber orientation.

Let us analyze whether our proposed PWMO satisfies the three criteria for well-designed local descriptors, which are defined in Section 1.

1. The two patch statistic attributes proposed in Eqs. (9) and (10) can inherit the robustness to illumination and noise variations as indicated in Section 2.1, thus PWMO can minish the intraperson variations caused by the illumination variations and noises.
2. The p-LBP/LXP encoding scheme and the spatial histogram model make the face representation less sensitive to the expression and pose variations and small registration error.
3. The combination of robust patch statistics, the p-LBP/LXP encoding scheme and the spatial histogram model can well capture the discriminative texture and structure information.
4. The computations of the two patch statistic attributes proposed in Eqs. (9) and (10) are efficient by taking advantage of integral image.

In conclusion, our proposed PWMO should be robust, discriminative and computationally efficient. Although the proposed local feature extraction approach (i.e., PWMO) can be directly applied to face representation, the dimensionality of the obtained feature sets is relatively high. Thus we should apply feature selection technology to reduce the feature dimensionality. Fisher linear discriminant (FLD) [5] proves to be a successful approach for discriminative feature selection of face images. However, in high dimensional space, FLD suffers heavily from 'small sample size (SSS)' problem, which means that the sample size is much smaller than the feature dimensionality. To address this problem, we adopt the block-based Fisher's linear discriminant (BFLD) approach proposed by Xie et al. [26]. Moreover, lots of work [13,24,25] have demonstrated that the features extracted from different facial areas take different discriminative information and thus should be assigned with different weights. We take advantage of the Fisher separation criterion (FSC) [25,35] to evaluate

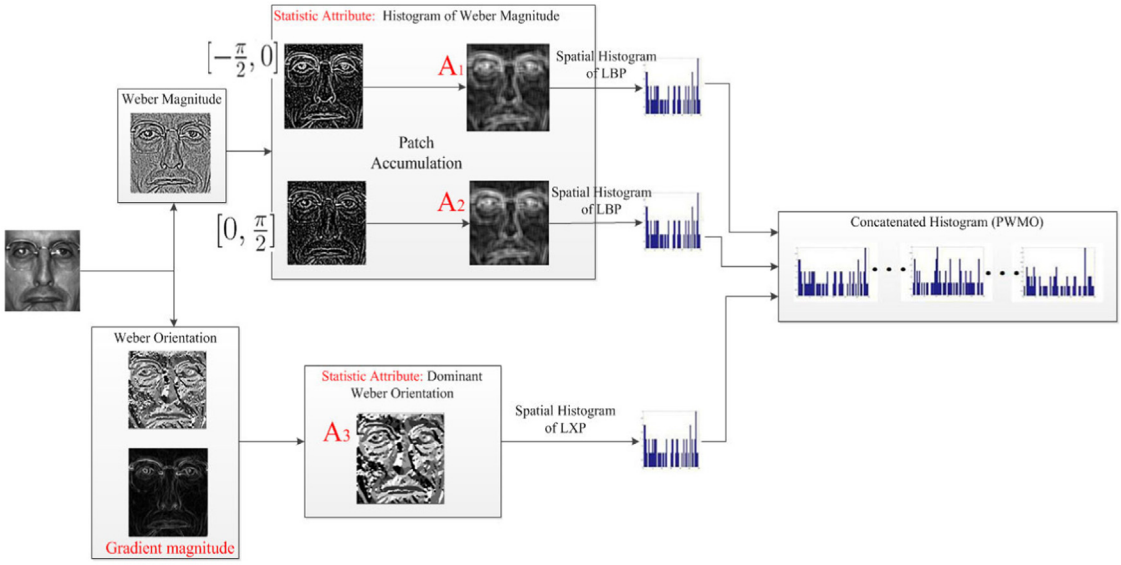


Fig. 7. An illustration of the framework based on PWMO for face representation.

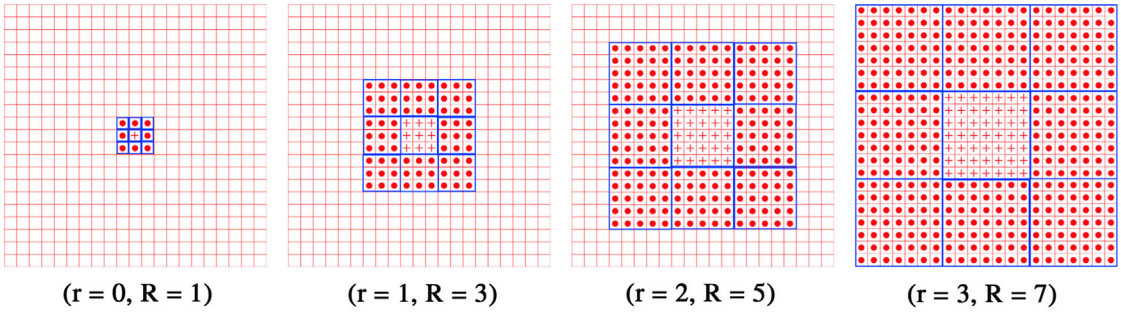


Fig. 8. Illustration of multi-scale PWMO for different (r,R).

the discriminative power of different block areas and then assign different weights to different blocks. The framework of PWMO feature extraction along with BFLD is illustrated in Algorithm 1.

Algorithm 1. Feature extraction using PWMO and block-based Fisher linear discriminant (FLD) matrices.

Input: An input face image, I; The learned block-based FLD projection matrix, U_i $(i=1,2,\ldots,M)$;
Output: M low-dimensional PWMO feature vectors, F_i;
1: For face image I, divide it into blocks B_i $(i=1,2,\ldots,M)$
2: **for** each block B_i, $i\in[1,M]$ **do**
3: Compute its Weber magnitude and orientation, as Eqs. (3)–(4)
4: Extract the PWM and PWO:
(i) Encode the Weber magnitude with p-LBP as Eqs. (6), (7), (9);
Encode the Weber orientation with p-LXP as Eqs. (6), (8), (10);
5: Partition PWM and PWO into nonoverlapping sub-blocks and calculate their histograms
6: Concatenate the spatial histograms of its sub-blocks together into a high-dimensional feature vector H_i
7: Calculate its low-dimensional vector F_i using the FLD transforms : $F_i=(U_i)^T H_i$.
8: **end for**

2.4. Single-scale vs. multi-scale

Multi-scale feature extraction can capture local salient patterns in different granularities, thus resulting in more powerfully discriminative ability. For example, multi-scale LBP, obtained by varying the sampling radius, has demonstrated better performance for both texture classification [14] and face recognition [34].

In this paper, we will show the performance of both single-scale PWMO (hereafter, PWMO for short) and multi-scale PWMO. Single-scale PWMO has lower feature dimensionality, while multi-scale PWMO could provide better performance. For single-scale PWMO, we empirically determine the optimal parameters, and note that the optimal (r,R) for PWM and PWO should not necessarily be equal. For multi-scale PWMO, we calculate the patch-based self-similarity by performing p-LBP and p-LXP at different scales of cells and blocks, i.e., different configurations of (r,R). In general, multi-scale PWMO can capture not only the microstructures but also the macrostructures, thus can improve the discriminative power compared with a single resolution of (r,R). An illustration of multi-scale PWMO with different (r,R) is given in Fig. 8. We can see that, for a small scale, the local micropatterns can be well represented, which is significant for discriminating facial details. On contrast, for a larger scale, the macropatterns, containing complementary information to small scale details, can be well captured.

3. Experimental results

In this section, extensive experiments are carried out for illustrating the effectiveness and efficiency of our proposed method. Specifically, two large publicly available face databases, FERET [36] and FRGC-204 [3], are utilized for performance evaluation of several state-of-the-arts' methods. These face databases and the following experiment setup include variant variations such as illumination, expression, aging, occlusion, noise, blur, and small pose.

We compare our proposed method with five other local descriptor-based methods: LBP [15], POEM [27], LGBP [24], LGXP [26] and spsLBP [20]. In order to keep more spatial information, in our experiments we empirically extract the LBP histograms in 8×8 blocks with 59 bins, each corresponding to a uniform pattern. The implementation of POEM is provided by Dr. Vu and the parameters are set according to [27]. The parameters of Gabor wavelets are set according to [26].

The nearest neighbor classifier is used throughout all experiments. For spatial histograms of all local descriptors, the similarity measure is defined by the histogram intersection:

$$d(H_1,H_2)=\sum_i \min(h_1^i,h_2^i), \tag{12}$$

where H_1 and H_2 are two histograms, and h_1^i, h_2^i are the i-th bin values, respectively.

For BFLD, we divide each face image into 16 (4×4) nonoverlapping blocks and perform BFLD block-wisely. We empirically retain 200 dimensions for each block, and thus the final dimensionality for a specific face image is 3200. Finally cosine distance is adopted as the similarity measure for each block.

3.1. Experiment I: FERET database

We use the standard FERET protocol [36] to conduct our experiments. The gallery set Fa consists of 1196 images of 1196 subjects. There are four probe sets: Fb (different expressions with gallery, 1195 images of 1196 subjects), Fc (different illumination conditions with gallery, 194 images of 194 subjects), Dup I (images taken later in time, 722 images of 243 subjects), Dup II (images taken at least 18 months after the corresponding gallery, 234 images of 75 subjects). All face images are properly aligned, cropped and resized to 128×128 with the centers of the eyes fixed at (29,34) and (99,34).[1] No further preprocessing is performed. Some sample face images from FERET are demonstrated in Fig. 9.

Parameters evaluation: For PWMO, it is not easy to theoretically determine the optimal parameter configuration. We test hundreds

[1] As indicated by [37], slight eye coordinates errors exist for some FERET images and the updated eye-coordinate files is provided by Dr. W. Deng.

Fig. 9. Sample face images in FERET face database, and the images in each row correspond to the same subject.

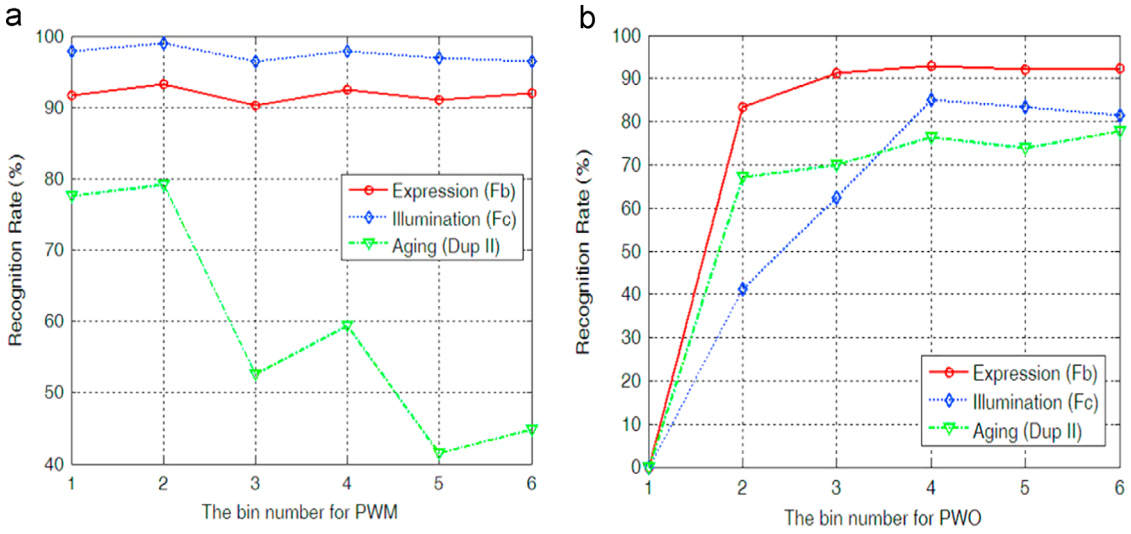


Fig. 10. The recognition rates versus the bin number for (a) PWM; (b) PWO.

of different parameter configurations $(r, R, bin_{wm}, bin_{wo})$ on FERET database and set the parameters empirically. Figs. 10 and 11 plot the recognition rates versus variant bin numbers (i.e., bin_{wm} and bin_{wo}) and (r,R) respectively, from which several useful conclusions can be drawn:

1. For PWM, bin_{wm} larger than 2 provides little help when expression (Fb) and illumination (Fc) variations exist, probably due to the unimodal distribution of Weber magnitude illustrated in Fig. 2. However, when aging (Dup II) variation exists, bin_{wm} larger than 2 largely degrades the performance, probably because that more bins are sensitive to the facial wrinkles in aging faces.
2. For PWO, as shown in Fig. 3, bin_{wo} fewer than 3 is not able to provide enough discriminative power, while the satisfying performance can be achieved by setting $bin_{wo} = 4$. Increasing bin number not only provides little help but also brings extra computational cost.
3. A slightly larger scale may extract more complementary and discriminative information and eliminate more noise, and thus achieve better performance. However, the scale cannot be too large. A much larger scale may fail to represent the local micropatterns and lead to worse robustness to illumination variants and thus achieve poorer performance.
4. With proper parameters, PWM is relatively more robust than PWO, especially under varying illumination conditions.

According to the above results, for single PWMO, we can regard the parameters with the higher accuracy as the optimal ones in terms of recognition rates. For PWM, we set $r=3$, $R=5$, $bin_{wm}=2$. For PWO, we set $r=1$, $R=3$, $bin_{wo}=4$. For multi-scale PWMO, We test hundreds of different configurations (r,R) on multiscale PWMO, and we obtain a relatively higher recognition rate when we set $(r,R)=\{(0,1),(1,3),(2,5),(3,7)\}$, $bin_{wm}=2$, $bin_{wo}=4$. This is the default parameter configuration in the following experiments.

Spatial histogram of local descriptors: Firstly, we want to check whether the spatial histogram of PWMO descriptor is suitable for face recognition, and conduct experiments on standard FERET protocol. The results are illustrated in Table 1, from which we could see that PWMO is a rather effective feature extraction approach. Especially, when illumination and aging variations exist, PWMO performs much better than LBP, spsLBP, POEM, LGBP, and is comparable to LGXP. It can also be concluded that (1) PWM and PWO can provide complementary discriminative information; and

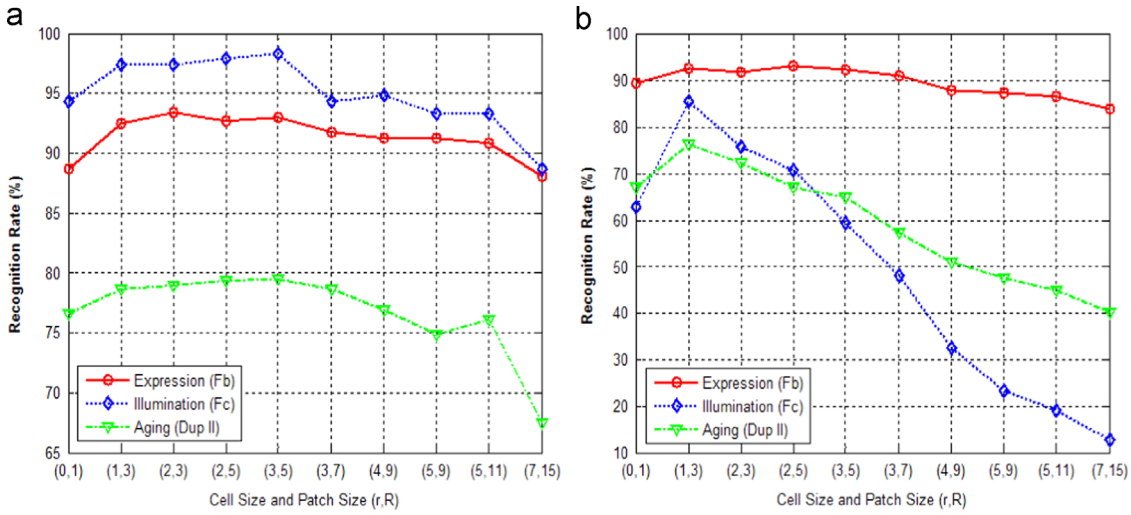


Fig. 11. The recognition rates versus the configuration of (r,R) for (a) PWM; (b) PWO.

Table 1
Recognition rates of different feature extraction methods on FERET database. Note the results with ∗ are directly cited from the original paper.

Methods	Fb (%)	Fc (%)	Dup I (%)	Dup II (%)	Average (%)
LBP [15]*	93	51	61	50	75
LGBP [24]*	96	96	74	70	87
LGXP [26]*	**98**	99	82	83	**92**
POEM [27]*	98	96	78	77	90
8-spsLBP[20]*	96	95	71	65	85
PWM	93	98	80	79	88
PWO	92	85	75	76	85
PWMO	94	99	82	81	90
Multi-scale PWMO	95	**99**	**84**	**83**	91

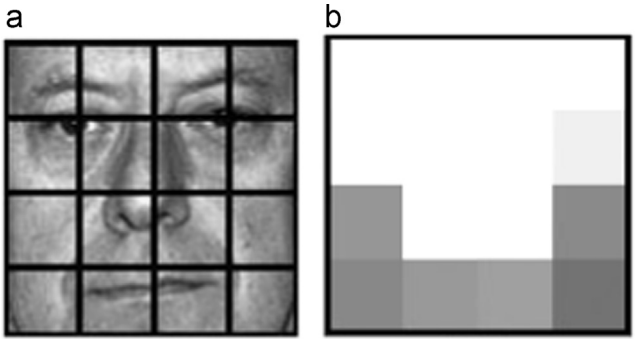


Fig. 12. Illustration of (a) the block partitioning and (b) weighting scheme based on Fisher separation criterion.

(2) multi-scale PWMO can further enhance the overall performance.

BFLD+block weighting scheme: To learn the BFLD projection matrix and calculate the weights for different blocks, we take advantage of the standard training set for FERET database, which consists of 1002 frontal images of 429 subjects.

The partitioning scheme is illustrated in Fig. 12(a), and the corresponding weights for different block areas are visualized in Fig. 12(b). As can be seen, the eye areas play a more important role in face recognition, which agrees with the conclusions in [15,13].

In order to help understand the performance of our proposed algorithms, we present a thorough analysis about the effectiveness of every important step and summarize into Table 2. From this table, we can see that the improvement mainly includes four important steps:

Table 2
The improvement of our proposed methods in terms of recognition rates on FERET database.

Methods	Fb (%)	Fc (%)	Dup I (%)	Dup II (%)
PWM	93	98	80	79
PWO	92	85	75	76
PWMO	94	99	82	81
Multi-Scale PWMO	95	99	84	83
PWMO+BFLD	98	99	93	92
PWMO+WBFLD	99	99	94	93
Multi-scale PWMO+BFLD	99	99	94	94
Multi-scale PWMO+WBFLD	**99**	**99**	**95**	**95**

Table 3
Performance comparison of proposed method with several state-of-the-art approaches on FERET database. Note the results with ∗ are directly cited from the original paper.

Methods	Fb (%)	Fc (%)	Dup I (%)	Dup II (%)
FERET97 Best [36]*	96	82	59	52
LBP [15]*	97	79	66	64
Fusing(Gabor and LBP) [17]*	98	98	90	85
POEM+WPCA [27]*	99	99	89	85
LBP+ESRC [37]*	97	95	94	92
(LGBP+LGXP)+BFLD [26]*	99	99	94	93
GOM [39]*	99	**100**	95	93
PWMO+BFLD	98	99	93	92
PWMO+WBFLD	99	99	94	93
Multi-scale PWMO+BFLD	99	99	94	94
Multi-scale PWMO+WBFLD	**99**	99	**95**	**95**

the combination of PWM and PWO, multi-scale scheme, BFLD and weighting scheme for BFLD. The combination of PWM and PWO helps PWMO obtain better performance than PWM and PWO upon all four subsets. The feature selection phase by BFLD improves a lot the recognition rates, even up to 11% on the Dup I and Dup II subsets. Furthermore, both multi-scale scheme and weighting scheme can further enhance the overall performance.

To further demonstrate the effectiveness of our proposed method, we compare it with other state-of-the-art methods reported in literatures. The results are summarized in Table 3. FERET97 Best [36] is the best identification performance of partially automatic algorithms (e.g., USC uses dynamic link architecture representation with elastic graph matching as similarity measure) tested in March 1997. Fusing

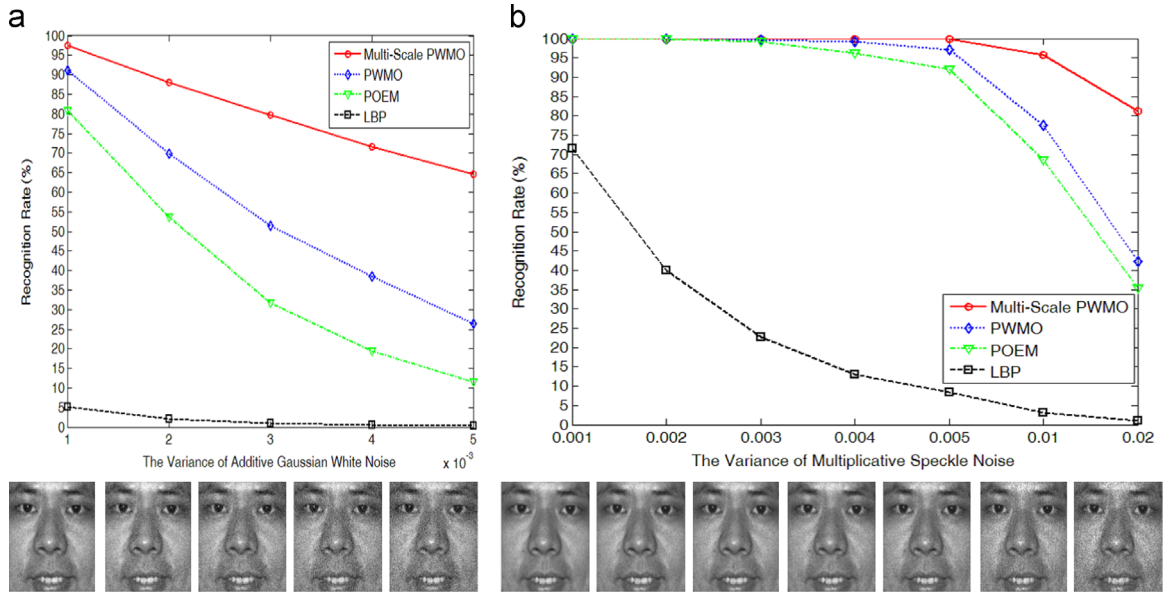


Fig. 13. (a) The recognition rates versus the variance of the additive Gaussian white noise; (b) the recognition rates versus the variance of the multiplicative speckle noise.

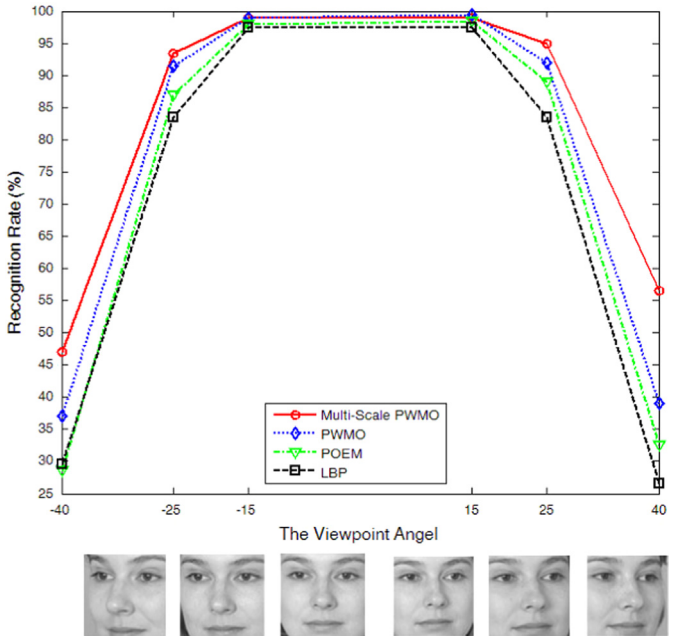


Fig. 14. The recognition rates versus the viewpoint angle.

(Gabor and LBP) [17] applies the feature-level fusion of the Gabor wavelets and LBP with PCA. POEM [27] represents patch-based self-similarity by the accumulated gradient magnitudes across different directions. LBP+ESRC [37] exploits LBP feature with the Extended Sparse Representation-Based Classifier (ESRC). (LGBP+LGXP)+BFLD [38] is the fusion of LGBP and LGXP. GOM [39] takes advantage of different kinds of ordinal measures on magnitude, phase, real, and imaginary components of Gabor images. As can be seen, multi-scale PWMO performs better than PWMO, and the combination of PWMO and BFLD can achieve high accuracy in all four probe sets. Moreover, the block weighting scheme (denoted as WBFLD) can further improve the performance. Our best result is comparable to those reported in the state-of-the-arts.

3.2. Experiment II: impact of noise and pose variations

Impact of noise: To our best knowledge, few works have been conducted to evaluate the tolerance of different face representations to noise. However, for the face recognition tasks under uncontrolled circumstances, noisy conditions are inevitable. Therefore, it is important for a local descriptor to be insensitive to noise for robust face recognition in real-world applications.

In this subsection, we test the robustness of PWMO to both additive and multiplicative noise, and also compare its performance with that of POEM and LBP. We conduct experiments on FERET data set, with Fa as the gallery and the noise-polluted (additive white Gaussian noise and multiplicative speckle noise) versions of Fa as the probes. The results are illustrated in Fig. 13. It is clear that LBP is vulnerable to noises and degrades much even with small noise level. Our proposed PWMO are much more robust than LBP and POEM under both the additive and multiplicative noisy conditions. Especially, exploiting multi-scale information tends to be more robust and discriminative in seriously noisy circumstances.

Impact of pose variations: Although many practical systems aim at the applications of frontal faces recognition, small pose variations are inevitable in real-world applications. We further empirically investigate the sensitiveness of different local descriptors to pose variations. We take advantage of the non-frontal subset of FERET database, which contains images captured at different view points from 200 subjects. We use the frontal images (ba) as the gallery, and the images taken at viewpoint angles of -40° (bh), -25° (bg), -15° (bf), $+15^\circ$ (be), $+25^\circ$ (bd), $+40^\circ$ (bc) are used as probes. The recognition rates versus viewpoint angles are illustrated in Fig. 14. We can see that all descriptors are robust to small pose variations ($\leq 25^\circ$), and the performance degrades much when the pose variation becomes larger. Although our proposed PWMO is more robust compared with LBP and POEM, pose-invariant face recognition is still an open problem [40].

3.3. Experiment III: FRGC-204

FRGC version 2.0 experiment 4 [3] is designed to compare different face verification technologies for face images with uncontrolled conditions, including illumination variations, expression changes, time elapse, blurred images and some occlusions. The data set for

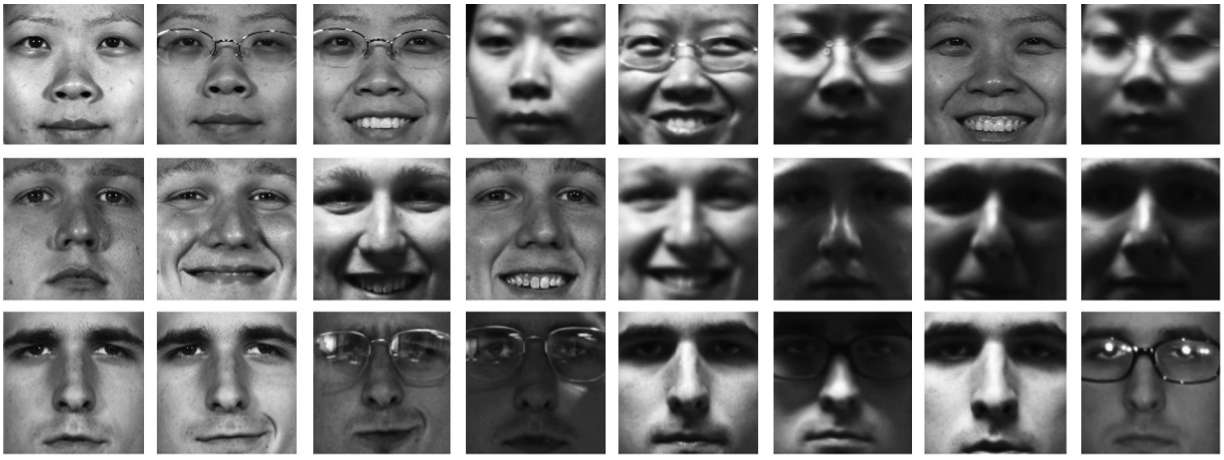


Fig. 15. Sample face images in FRGC face database, and the images in each row correspond to the same subject.

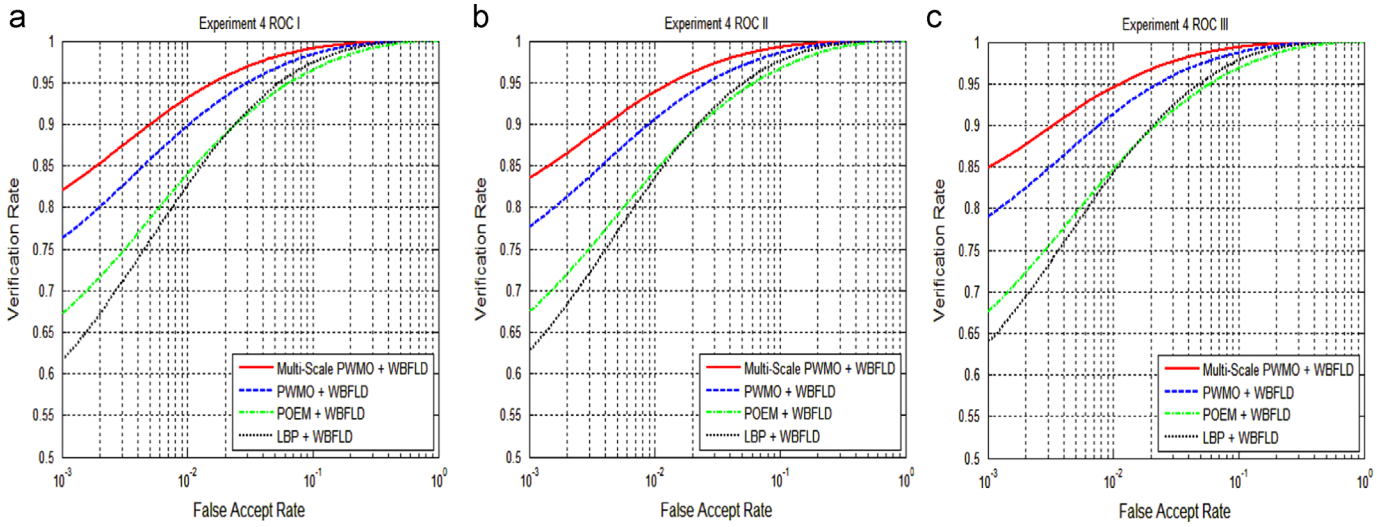


Fig. 16. ROC curves corresponding to the LBP, POEM, PWMO and multi-scale PWMO based face representation approaches on FRGC2.0 Experiment 4: (a) ROC I. (b) ROC II. (c) ROC III.

FRGC-204 .is divided into training and validation partitions. The training partition was collected from 222 subjects in the 2002–2003 academic year and consists of 12,776 images (6388 controlled still image and 6388 uncontrolled still image). The validation partition was collected from 466 subjects in the 2003–2004 academic year, which consists of a target set with 16,028 controlled images and a query set with 8014 uncontrolled images. Therefore, FRGC-204 is rather challenging due to its large scale and variant unexpected variations. In our experiments, all the face images are aligned, cropped and resized via the provided eye coordinates in the same manner with that of Section 3.1 Some sample face images from FRGC are demonstrated in Fig. 15.

According to the protocol of FRGC, three receiver operator characteristic (ROC) curves are generated: ROC I, ROC II, ROC III, each corresponding to the images collected within semesters, within a year and between semesters, respectively. Typically, for ease of performance comparison, most researchers report the verification rates (VR) at 0.1% false acceptance rate (FAR) in their work.

In our experiments, the training partition is utilized to learn the BFLD projection matrix and calculate the weights for different blocks. Fig. 16 plots the three ROC curves of different local descriptors, all with WBFLD or FLD as the feature selector. Comparisons with BEE (Biometric Experimentation Environment) baseline and several representative methods are summarized in Table 4. We divide the methods into three categories according to the features exploited in the methods: global feature, local feature and ensemble of global and local feature. Global feature based on the subspace learning (e.g., PCA) or spatial-frequency transformation technologies (e.g., Fourier feature) to explore the whole face image. Local feature extracts and encodes the local structure information of facial regions in a manner of patch filtering. Ensemble of global and local feature is a combination of the global and local feature. As can be seen, our proposed methods are much better than the BEE baseline and can provide better or comparable results than most of the methods in the literature. Our proposed method outperforms most competing local descriptors, including LBP, POEM, LGBP, and LGXP. The VR of 84.9% for ROC III achieved by multi-scale PWMO is comparable to that of the fusion method of [26] where LGBP and LGXP are combined using score-level combination. The results in [17], based on the kernel combination of Gabor and LBP, are worse than ours. However, our methods still cannot outperform some leading methods such as GOM [39] and the ensemble of the global Fourier feature and local Gabor feature [42]. An important conclusion drawn from Table 4 is that the combination of global and local features can further boost the performance of face recognition systems.

3.4. *Efficiency analysis*

In this subsection, we discuss the efficiency of PWMO by comparing the storage requirement and the computational cost, which are

Table 4
Verification rates (%) for several recently published state-of-the-art face representation approaches on FRGC 2.0 Experiment 4 when FAR=0.1%. Note the results with * are directly cited from the original paper.

Methods	ROC I (%)	ROC II (%)	ROC III (%)
Global feature			
BEE Baseline (PCA) [3]*	16.08	15.18	14.01
Hybrid Fourier Feature [10]*	81.82	81.50	81.14
Local feature			
LBP [15]	61.6	62.7	63.8
POEM [27]	67.1	67.4	67.6
Gabor [41]*	–	–	76
Fusion of LBP and Gabor [17]*	–	–	83.6
LGBP [26]*	79.6	80.3	81.0
LGXP [26]*	78.2	78.6	78.9
Fusion of LGBP and LGXP [26]*	83.6	84.3	84.9
GOM (one stage) [39]*	85.6	85.2	84.8
GOM (two stages) [39]*	86.2	85.7	85.2
PWMO	76.3	77.6	79.0
Multi-Scale PWMO	81.99	83.44	84.87
Ensemble of global and local feature			
Fourier and Gabor feature [42]	–	–	89

Table 5
Complexity comparison for different local feature extraction methods. Note the results with * are directly cited from the original paper.

Methods	Dimension	Time per image (ms)
LBP [15]	15,104	12.61
POEM [27]	45,312	37.97
LGBP [24]	604,160	1052.84
LGXP [26]	604,160	1065.80
Fusing(LGBP+LGXP) [26]	1,208,320	2126.46
GOM [39]*	1,474,560	700*
PWMO	45,312	50.87
Multi-Scale PWMO	181,248	202.84

critical for large-scale face recognition applications, with other competing local feature extraction algorithms. We measure the storage cost by the feature dimensionality per face image, and measure the computational cost by the average processing time per image calculated from 1000 face images (128×128). The computational cost of GOM method are referred to the original paper. All the rest of the algorithms are implemented using 64-bit Matlab platform on a PC with Intel I7 CPU (2.80 GHz) and 4 GB memory. The results are summarized in Table 5.

As can be seen from Tables 3 to 5, our proposed PWMO/multi-scale PWMO is a good balance between efficiency and discriminative power. More specifically, our proposed method is much more discriminative than LBP and the recently proposed POEM, and can achieve comparable results to the fusion of LGBP and LGXP with 1/7 storage requirement and 1/10 computational cost, which implies the superiority of PWMO for real-time face recognition tasks.

4. Conclusion

In this paper we have proposed a novel local descriptor, called patterns of Weber magnitude and orientation (PWMO), for face representation and recognition. The effectiveness of PWMO comes from several aspects including the insensitiveness of Weber magnitude and orientation to variant variations such as illumination and noise, the patch-based self-similarity encoding scheme, the spatial histogram modeling, and the multi-resolution representation. To further reduce the feature dimensionality, we adopted BFLD to select the most discriminative feature sets, and the block weighting scheme based on Fisher separation criteria has been incorporated for discriminant classification. Experimental results on two publicly available face databases have evidently illustrated the effectiveness and efficiency of the proposed method. Especially on the challenging FRGC-204 face database, we have achieved comparable verification performance as the recently proposed method based on the fusion of LGBP and LGXP, while requiring much less storage and computational cost. Due to its remarkable performance and low computational cost in face recognition applications, we expect that the proposed method is a good choice for the recognition of other objects.

Acknowledgments

We thank Dr. W. Deng for providing the updated eye coordinates for the FERET database. Special thanks are given to Dr. N.S. Vu for sharing his implementation of the POEM descriptor.

References

[1] W. Zhao, R. Chellappa, A. Rosenfeld, P. Phillips, Face recognition: a literature survey, ACM Comput. Surv. (2003) 399–458.
[2] S. Li, A. Jain, Handbook of Face Recognition, 2nd edition, Springer-Verlag, London, 2011.
[3] P. Phillips, P. Flynn, T. Scruggs, K. Bowyer, C. Jin, K. Hoffman, J. Marques, M. Jaesik, W. Worek, Overview of the face recognition grand challenge, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005, pp. 947–954.
[4] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cogn. Neurosci. 3 (1991) 71–86.
[5] P. Belhumeur, J. Hespanha, D. Kriegman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell. 19 (1997) 711–720.
[6] K. Kwak, W. Pedrycz, Face recognition using an enhanced independent component analysis approach, IEEE Trans. Neural Netw. 18 (2007) 530–541.
[7] X. He, S. Yan, Y. Hu, P. Niyogi, H. Zhang, Face recognition using Laplacianfaces, IEEE Trans. Pattern Anal. Mach. Intell. 27 (2005) 328–340.
[8] J.Y.J. Yang, D. Zhang, B. Niu, Globally maximizing, locally minimizing: unsupervised discriminant projection with applications to face and palm biometrics, IEEE Trans. Pattern Anal. Mach. Intell. 29 (2007) 650–664.
[9] M. Yang, Kernel eigenfaces vs. kernel Fisherfaces: face recognition using kernel methods, in: International Conference on Automatic Face and Gesture Recognition, 2002, pp. 215–220.
[10] W. Hwang, H. Wang, H. Kim, S. Kee, J. Kim, Face recognition system using multiple face model of hybrid fourier feature under uncontrolled illumination variation, IEEE Trans. Image Process. 20 (2011) 1152–1165.
[11] C. Liu, D. Dai, Face recognition using dual-tree complex wavelet features, IEEE Trans. Image Process. 18 (2009) 2593–2599.
[12] Z. Hafed, M. Levine, Face recognition using the discrete cosine transform, Int. J. Comput. Vis. 43 (2001) 167–188.
[13] J. Zou, Q. Ji, G. Nagy, A comparative study of local matching approach for face recognition, IEEE Trans. Image Process. 16 (2007) 2617–2628.
[14] T. Ojala, M. Pietikäinen, T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell. 24 (2002) 971–987.
[15] T. Ahonen, A. Hadid, M. Pietikäinen, Face description with local binary patterns: application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (2006) 2037–2041.
[16] D. Huang, C. Shan, M. Ardabilian, Y. Wang, L. Chen, Local binary patterns and its applications to facial image analysis: a survey, IEEE Trans. Syst. Man Cybern.—Part C: Appl. Rev. 41 (2011) 765–778.
[17] X. Tan, B. Triggs, Fusing gabor and lbp feature set for kernel-based face recognition, in: Proceedings of the IEEE International on Workshop on Analysis and Modeling of Faces and Gestures, 2007, pp. 235–249.
[18] B. Zhang, Y. Gao, S. Zhao, J. Liu, Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor, IEEE Trans. Image Process. 19 (2010) 533–544.
[19] L. Wolf, T. Hassner, T. Taigman, Effective unconstrained face recognition by combining multiple descriptors and learned background statistics, IEEE Trans. Pattern Anal. Mach. Intell. 33 (2011) 1978–1990.
[20] Z. Chai, H. Mendez-Vazquez, R. He, Z. Sun, T. Tan, Semantic pixel sets based local binary patterns for face recognition, in: Proceedings of the ACCV, Springer, Berlin Heidelberg, 2013, pp. 639–651.
[21] J. Daugman, Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, J. Opt. Soc. Am. 2 (1985) 1160–1169.
[22] L. Wiskott, J. Fellous, N. Kruger, C. Malsburg, Face recognition by elastic bunch graph matching, IEEE Trans. Pattern Anal. Mach. Intell. 19 (1997) 775–779.

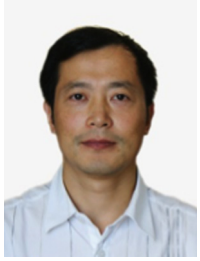
[23] W. Deng, J. Hu, J. Guo, Gabor-eigen-whiten-cosine: a robust scheme for face recognition, in: AMFG, 2005, pp. 336–349.
[24] W. Zhang, S. Shan, W. Gao, X. Chen, H. Zhang, Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition, in: Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2005, pp. 786–791.
[25] B. Zhang, S. Shan, X. Chen, W. Gao, Histogram of Gabor phase patterns (hgpp): a novel object representation approach for face recognition, IEEE Trans. Image Process. 16 (2006) 57–68.
[26] S. Xie, S. Shan, X. Chen, J. Chen, Fusing local patterns of Gabor magnitude and phase for face recognition, IEEE Trans. Image Process. 19 (2010) 1349–1361.
[27] N. Vu, A. Caplier, Enhanced patterns of oriented edge magnitudes for face recognition and image matching, IEEE Trans. Image Process. 21 (2011) 1352–1365.
[28] P.H. Gregson, Using angular dispersion of gradient direction for detecting edge ribbons, IEEE Trans. Pattern Anal. Mach. Intell. 15 (1993) 682–696.
[29] B. Heisele, P. Ho, J. Wu, T. Poggio, Face recognition: component-based versus global approaches, Comput. Vis. Image Underst. 91 (2003) 6–12.
[30] B. Wang, W. Li, Z. Li, Q. Liao, Patterns of weber magnitude and orientation for face recognition, in: IEEE International Conference on Image Processing (ICIP), 2012, pp. 1441–1444.
[31] J. Chen, S. Shan, C. He, G. Zhao, M. Pietikäinen, X. Chen, W. Gao, Wld: a robust local image descriptor, IEEE Trans. Pattern Anal. Mach. Intell. 32 (2010) 1705–1720.
[32] B. Wang, W. Li, W. Yang, Q. Liao, Illumination normalization based on Weber's law with application to face recognition, IEEE Signal Process. Lett. 18 (2011) 462–465.
[33] T. Zhang, Y. Tang, B. Fang, Z. Shang, X. Liu, Face recognition under varying illumination using gradient faces, IEEE Trans. Image Process. 18 (2009) 2599–2606.
[34] C. Chan, J. Kittler, K. Messer, Multi-scale local binary pattern histogram for face recognition, Lecture Notes in Computer Science, vol. 4642, 2007, pp. 809–818.
[35] R. Parry, I. Essa, Feature weighting for segmentation, in: Proceedings of the 5th International Conference on Music Information Retrieval, 2004, pp. 116–119.
[36] P.J. Phillips, H. Moon, P. Rizvi, P. Rauss, The Feret evaluation methodology for face recognition algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 22 (2000) 1090–1104.
[37] W. Deng, J. Hu, J. Guo, Extended src: undersampled face recognition via intra-class variant dictionary, IEEE Trans. Pattern Anal. Mach. Intell. 34 (2012) 1864–1870.
[38] S. Xie, S. Shan, X. Chen, J. Chen, Fusing local patterns of Gabor magnitude and phase for face recognition, IEEE Trans. Image Process. 19 (2010) 1349–1361.
[39] Z. Chai, Z. Sun, H. Mendez-Vazquez, R. He, T. Tan, Gabor ordinal measures for face recognition, IEEE Trans. Inf. Forensics Secur. 9 (2014) 14–26.
[40] X. Zhang, Y. Gao, Face recognition across pose: a review, Pattern Recognit. 42 (2009) 2876–2896.
[41] C. Liu, Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance, IEEE Trans. Pattern Anal. Mach. Intell. 28 (2006) 725–737.
[42] Y. Su, S. Shan, X. Chen, W. Gao, Hierarchical ensemble of global and local classifiers for face recognition, IEEE Trans. Image Process. 18 (2009) 1885–1896.

Yinyan Jiang received the B.E degree in Electronic Engineering from Tsinghua University, Beijing, China, in 2012. She is currently a master student working on face recognition in Tsinghua University. Her research interests include image processing and pattern recognition in biometrics.

Biao Wang was born in Huanggang, HuBei, China, in 1986. He received the B.E. degree from Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan, China, in 2008, and the Ph.D. degree in Electronic Engineering from Tsinghua University, Beijing, China, in 2013. Since July 2013, he serves as a senior researcher in the Research & Development Center, Samsung Electronics, Beijing, China. His research interests include applications of image processing and pattern recognition in biometrics.

Yicong Zhou received his B.S. degree from Hunan University, Changsha, China, and his M.S. and Ph.D. degrees from Tufts University, Massachusetts, USA, all degrees in electrical engineering. He is currently an Assistant Professor in the Department of Computer and Information Science at University of Macau, Macau, China. His research interests focus on multimedia security, image/signal processing, pattern recognition and medical imaging. Dr. Zhou is a member of the IEEE and SPIE (International Society for Photo-Optical Instrumentations Engineers).

Weifeng Li received the M.E. and Ph.D. degrees in Information Electronics at Nagoya University, Japan, in 2003 and 2006, respectively. He joined the Idiap Research Institute, Switzerland in 2006, and in 2008 he moved to Swiss Federal Institute of Technology, Lausanne (EPFL), Switzerland, as a Research Scientist. Since 2010 he has been an Associate Professor in the Department of Electronic Engineering/Graduate School at Shenzhen, Tsinghua University, China. His research interests lie in the areas of audio and visual signal processing, biometrics, human–computer interactions (HCI), and machine learning techniques. He is a member of the IEEE and IEICE.



Qingmin Liao received the B.S. degree in radio technology from the University of Electronic Science and Technology of China, Chengdu, China, in 1984, and the M.S. and Ph.D. degrees in signal processing and telecommunications from the University of Rennes 1, Rennes, France, in 1990 and 1994, respectively. Since 1995, he has been joining with Tsinghua University, Beijing, China. In 2002, he became Professor in the Department of Electronic Engineering of Tsinghua University. Since 2010, he has been the Director of the Division of Information Science and Technology in the Graduate School at Shenzhen, Tsinghua University. He is also affiliated with the Shenzhen Key Laboratory of Information Science and Technology (Director), China. Over the last 30 years, he has published over 100 peer-reviewed journal and conference papers. His research interests include image/video processing, transmission and analysis; biometrics; and their applications to teledetection, medicine, industry, and sports.