Neurocomputing 120 (2013) 325–335

Contents lists available at ScienceDirect

Neurocomputing

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

WLBP: Weber local binary pattern for local image description

Fan Liu, Zhenmin Tang, Jinhui Tang*

School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, PR China

ARTICLE INFO

Article history:
Received 14 November 2011
Received in revised form
25 May 2012
Accepted 19 June 2012
Available online 30 March 2013

Keywords:
Local binary pattern
Weber's law
Differential excitation
Face recognition
Texture classification

ABSTRACT

In this paper, we propose a local descriptor, called Weber Local Binary Pattern (WLBP[1]), which effectively combines the advantages of WLD and LBP. Specifically, WLBP consists of two components: differential excitation and LBP. The differential excitation extracts perception features by Weber's law, while the LBP (Local Binary Pattern) can describe local features splendidly. By computing the two components, we obtain wo images: differential excitation image and LBP image, from which a WLBP histogram is constructed. The differential excitation was extended by bringing in Laplacian of Gaussian (LoG), which makes WLBP robust to noise. By designing a new quantization method, the discriminabilty of WLBP was enhanced. The proposed method is evaluated on the face recognition problem under different challenges. Experimental results show that WLBP performs better than WLD and LBP. Meanwhile, it is robust to time, facial expressions, lightings, pose and noise. We also conduct experiments on Brodatz and KTH-TIPS2-a texture databases, which demonstrate that WLBP is a powerful texture descriptor.

1. Introduction

Nowadays, local descriptors, which are distinctive and invariant to many kinds of geometric and photometric transformations, have been gaining more and more attention because of their promising performance [22]. And a large variety of methods using local features have been developed. Existing descriptors can be categorized into two classes: one is a sparse descriptor, which first detects the interest points in a given image, and then samples a local patch and describes its invariant features [23,24]; the other is a dense descriptor, which extracts local features pixel by pixel over the input image [2]. Scale invariant feature transform (SIFT) [25] and histogram of oriented gradients (HOG) [26] are typical sparse descriptors, while among the most popular dense descriptors are the Gabor wavelet [32] and local binary pattern (LBP) [1]. They are being widely utilized in a large number of applications, such as baseline matching [20,21], human detection [26–28], face recognition [11,29,30], object recognition [25], and texture classification [1,31].

Recently, Chen et al. [2] proposed a simple and powerful local descriptor, Weber Local Descriptor (WLD), which was inspired by a psychological law, Weber's law [3]. Since most images are eventually to be observed and interpreted by humans, an ideal image descriptor has to take into account the effects of human psychology and psychophysics. Specially, WLD consists of two components: differential excitation and orientation. The differential excitation component is a function of the ratio between two terms: one is the relative intensity difference of a current pixel and its neighbors; the other is the intensity of the current pixel. The orientation is the gradient orientation of the current pixel [2]. Experiments show that WLD impressively outperforms the other widely used descriptors (e.g., Gabor and SIFT) [2]. However, according to the definition of WLD, the orientation component only considers the four pixels on the horizontal and vertical directions, which cannot fully represent local information. In addition, the orientation and differential excitations are just averagely quantized into some intervals. This kind of simple quantization method not only leads to information loss further, but also limits the performance of WLD.

In this paper, we propose a new local descriptor, Weber Local Binary Pattern (WLBP). Similarly, WLBP also consists of two components: differential excitation and LBP [1]. For the original differential excitation component, we can find that the operator calculating the intensity differences of a current pixel against its neighbors is actually Laplacian operator. Since Laplacian operator is sensitive to noise, we replace it with Laplacian of Gaussian (LoG) in our method. As a well-known local descriptor, LBP has been proven to be highly discriminative, computationally efficient and invariant to monotonic gray level changes. It does better than the orientation component of WLD in extracting local features. With differential excitation component, we extract local salient patterns on human perception. By LBP component, local micro-patterns corresponding to bright/dark spots, edges and flat areas etc. are computed. For an input image, we compute two components of WLBP. By building statistics on LBP patterns along with the local perception patterns of differential excitation, we represent an

* Corresponding author. Tel.: +86 25 8431 7297; fax: +86 25 8431 5510.
E-mail addresses: fanfanliu.njust@gmail.com (F. Liu), tzm.cs@mail.njust.edu.cn (Z. Tang), tangjh1981@acm.org (J. Tang).

[1] WLBP is the abbreviation of Weber Local Binary Pattern.

0925-2312/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.neucom.2012.06.061

input image (or image region) with a histogram, which we call a WLBP histogram.

The rest of this paper is organized as follows: in Section 2, we describe the details of the proposed local descriptor WLBP. In Section 3, we carry out experiments on face image databases to demonstrate the effectiveness and robustness of WLBP. In Section 4, we conduct two experiments of texture classification. Finally, conclusions are presented in Section 5.

2. Weber local binary pattern

In this section, we first give a brief description of Weber's law. Then we detail the new differential excitation and LBP. After that we describe the proposed WLBP descriptor and present how to compute a WLBP histogram for an input image (or image region). Finally, we analyze the advantages of WLBP by comparing WLBP with WLD and LBP.

2.1. Weber's law

Most of us can easily catch a whispered voice in a quiet room, but in a noisy environment we may not hear someone shouting near our ear. This is the essence of Weber's law, observed by German physiologist Weber (1795–1878). It states that the ratio of the smallest perceptual increment in a stimulus to the background stimulus intensity is a constant [3]. This relationship, which is known as Weber's law, is shown as the following equation:

$$\frac{\Delta I}{I} = K \tag{1}$$

where ΔI represents the increment threshold (just noticeable difference for discrimination); I represents the initial stimulus intensity and K is a const proportion. The fraction $\Delta I/I$ is known as the *Weber fraction*.

Weber's law reveals the relationship between mental quantity and physical quantity. It, more simply stated, says that the stimulus change (ΔI) can be just noticeable only when the *Weber fraction* $\Delta I/I$is a constant. Therefore, when the change of a stimulus (such as sound, lighting) is smaller than this constant ratio of the original stimulus, it would be recognized as a background noise rather than a valid signal. From Eq. (1), we can find that Weber's law is useful in image processing [4]. Bruni and Vitulano used this law for scratch detection on digital film materials [18]. Phiasai et al. employed a Weber ratio to control the strength of watermark [19]. Moreover, the differential excitation component of WLD also derives from Weber's law.

2.2. Differential excitation

For WLD, the differential excitation $\xi(x_c)$ of a current pixel x_c is calculated like this

$$\Delta I = \sum_{i=0}^{p-1} (\Delta x_i) = \sum_{i=0}^{p-1} (x_i - x_c) \tag{2}$$

$$G_{ratio}(x_c) = \frac{\Delta I}{I} \tag{3}$$

$$\xi(x_c) = \arctan[G_{ratio}(x_c)] = \arctan\left[\frac{\Delta I}{I}\right] = \arctan\left[\frac{\sum_{i=0}^{p-1}(x_i - x_c)}{x_c}\right] \tag{4}$$

Here x_i is the i-th neighbor pixel of current pixelx_c, I equals current pixel x_c, and ΔI is the intensity differences of current pixel against its neighbors. Following the hints in Weber's law, we can find that Eq. (3) actually derives from Eq. (1). Where the arctangent function in Eq. (4) is used to prevent the output from being too large and thus could partially suppress the side-effect of noise. Note that the range of $\xi(x_c)$is$[-\pi/2, \pi/2]$. If $\xi(x_c)$is positive, it means that the surroundings are lighter than current pixel. In contrast, if $\xi(x_c)$ is negative, it means that the surroundings are darker than current pixel. For WLD, ξ is divided into several equal bins to build histograms.

In our approach, the differential excitation is different from that of WLD. We optimize the design of the differential excitation operator. As can be seen from Eq. (2), it is actually Laplacian operator, which is second-order derivative. Therefore, the formula (2) can be expressed as

$$\Delta I = \nabla^2 = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{5}$$

Where $f(x,y)$ represents the image and (x,y) is the position of current pixel x_c. ∇^2 is the second-order derivative. Since Laplacian operator is very sensitive to noise, we utilize Laplacian of Gaussian (LoG) to compute ΔI in our method, which is described as follows:

$$h(x,y) = \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) \tag{6}$$

$$g(x,y) = h(x,y) \otimes f(x,y) \tag{7}$$

$$\Delta I = \nabla^2 g = \nabla^2[h(x,y)\otimes f(x,y)] = \frac{1}{\pi\sigma^4}\left(\frac{x^2+y^2}{2\sigma^2} - 1\right)\exp\left(-\frac{1}{2\sigma^2}(x^2+y^2)\right) \tag{8}$$

where $h(x,y)$ is a 2D Gaussian function, $g(x,y)$ is the result of convolution of $h(x,y)$ and $f(x,y)$, ∇^2 is the second-order derivative. Then, we compute the ratio of the differences ΔI of Eq. (8) to the intensity of current pixelx_c($x_c = f(x,y) = I$), as is described by the following formula:

$$G_{ratio}(x_c) = \frac{\Delta I}{I} = \frac{\nabla^2 g}{I} \tag{9}$$

After that, the differential excitation of the current pixel x_c is computed as

$$\xi(x_c) = \arctan[G_{ratio}(x_c)] = \arctan\left[\frac{\Delta I}{I}\right] = \arctan\left[\frac{\nabla^2 g}{I}\right] \tag{10}$$

For the quantization of $\xi(x_c)$, we propose a new method, which can improve the representation ability of the differential excitation operator. Due to the fact that *Weber fraction* $\Delta I/I$ is a constant, we set a constant K to segment the range of ξ into two classes: high perception pattern and low perception pattern. Then, the interval $[-K, K]$ is classified into the low perception interval. $[-\pi/2, -K]$ and $[K, \pi/2]$ are high perception intervals. As high perception patterns include a majority of detail information, we divide the high perception intervals further. We assume that ξ is segmented into S intervals. According to symmetry, S is odd. Then S intervals are shown in the following formula:

$$l_s = \begin{cases} \left[-\left(m\left(\frac{\pi-2K}{S-1}\right)+K\right), -\left((m-1)\left(\frac{\pi-2K}{S-1}\right)+K\right)\right] \\ [-K, K] \\ \left[(m-1)\left(\frac{\pi-2K}{S-1}\right)+K, m\left(\frac{\pi-2K}{S-1}\right)+K\right] \end{cases} \quad (m = 1, \ldots (S-1)/2, s = 1, \ldots S) \tag{11}$$

2.3. Local binary pattern (LBP)

LBP operator, proposed by Ojala et al. [1], is a powerful means of texture description. It generates a binary number by thresholding each pixel of the neighborhood $\{g_0, \ldots g_{p-1}\}$ with the central pixel "g_c". The LBP value of the center pixel is obtained by converting the binary number into a decimal one. Using circular neighborhoods and linearly interpolating the pixel values allows the choice of any radius R, and number of pixels in the neighborhood P, to form an (P, R) neighborhood. Function (12) gives the computation

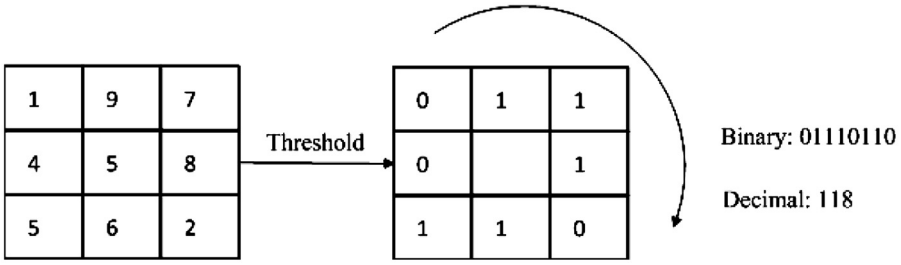


Fig. 1. Example of computing $LBP_{8,1}$.

of $LBP_{P,R}$:

$$LBP_{P,R}=\sum_{i=0}^{P-1}s(g_i-g_c)2^i$$
$$s(x)=\begin{cases}1, x>0\\ 0, x\le 0\end{cases} \quad (12)$$

Pixels which are greater than the central pixel are mapped to 1, otherwise 0. See Fig. 1 for an illustration of computing LBP feature.

Generally, $LBP_{P,R}$ produces 2^P different values. A subset of these 2^P binary patterns is called uniform patterns. The uniform pattern is a binary pattern which contains at most two bitwise transitions from 0 to 1 or vice versa. The uniformity measure, $U(x)$, presented in Eq. (12) records the number of bitwise transitions from 0 to 1 or vice versa. The uniform pattern contains in total $(P-1)P+2$ binary patterns. Other patterns, where $U(x)>2$, are regarded as non-uniform patterns. The uniform LBP operator, $LBP^{u2}_{P,R}$, is defined as

$$LBP^{u2}_{P,R}(x,y)=\begin{cases}I(LBP_{P,R}(x,y)) & \text{if } U(LBP_{P,R})\le 2,\\ & I(z)\in[0,(P-1)P+1]\\ (P-1)P+2 & \text{otherwise}\end{cases} \quad (13)$$

where

$$U(LBP_{P,R})=\sum_{p=1}^{P}|s(g_p-g_c)-s(g_{p-1}-g_c)|+|s(g_{p-1}-g_c)-s(g_0-g_c)|$$

Superscript $u2$ shown in Eq. (12) indicates that the uniform pattern is with a U value of at most 2. If $U(x)$ is smaller than 2, the current pixel will be labeled by an index function $I(z)$. Otherwise, it will be labeled as $(P-1)P+2$. By the index function $I(z)$, each of the uniform patterns is assigned to a particular index.

Ojala noticed that in their experiments with texture images, uniform patterns account for a bit less than 90% of all patterns when using the (8,1) neighborhood and for 70% in (16,2) neighborhood. Uniform pattern can be regarded as the basic properties of image, because most patterns of image are uniform pattern. Therefore, the "uniform" LBP was selected in our method.

2.4. WLBP descriptor

As one of the variables for many local descriptors [4], histogram is often utilized to express descriptors. In this paper, we also utilize histogram to express WLBP descriptor.

WLBP consists of two components: differential excitation and LBP, which are both described above. As shown in Fig. 2, for a given image, we compute the pattern value for every pixel by using the uniform LBP operator $LBP^{u2}_{P,R}$. And the differential excitation of every pixel is computed by formula (10). Then, we get two images: differential excitation image and LBP image. According to these images, we first construct the 2D histogram $\{WLBP(s,t)\}$, $(s=1,\ldots,S, t=1,\ldots,T)$ of the original image. Note that the size of this 2D histogram is $T\times S$, where S is the number of intervals of ξ, T is the total number of LBP's patterns. In other words, in this 2D histogram, each column corresponds to a pattern t of LBP, and each row corresponds to a differential excitation interval. Thus, the value of each cell $WLBP(s,t)$ corresponds to the frequency of the certain differential excitation interval l_s and the LBP pattern t.

To enhance the discriminability, the 2D histogram $\{WLBP(s,t)\}$ is further encoded into 1D histogram H. We use each row of 2D histogram to form a 1D histogram $H(s),(s=1,\ldots,S)$. Each sub-histogram $H(s)$ corresponds to the differential excitation interval l_s. Concatenating the S sub-histograms, we obtain the 1D histogram $H=\{H_s\}$, $s=1,\ldots,S$. In our approach, according to the equation of $T=(P-1)*P+3$, T equals to 59 as the values of the two subscripts P and R in the uniform LBP operator $LBP^{u2}_{P,R}$ are assigned with 8 and 2 respectively. While the parameters S (intervals number of ξ) and K(threshold of Eq. (11)) can be set freely according to the real condition, which will be discussed in Section 3.3.3.

2.5. Comparison with WLD and LBP

The reason we select LBP as a component of WLBP is that LBP can extract more local structure information than Orientation. In addition, LBP has proven to be highly discriminative. Take Fig. 3 for example, the formula for computing LBP is shown in the following equation:

$$LBP_{P,R}=\sum_{i=0}^{P-1}s(g_i-g_c)2^i$$
$$s(x)=\begin{cases}1, x>0\\ 0, x\le 0\end{cases} \quad (14)$$

While the Orientation of WLD is computed by the following equation:

$$\theta(x_c)=\arctan\left(\frac{x_7-x_3}{x_5-x_1}\right) \quad (15)$$

From (14) and (15), we can find that LBP makes full use of the nine pixel while Orientation operator only utilizes four pixels (x_1, x_3, x_5, x_7). That is to say, if x_1, x_3, x_5, x_7 are given, $\theta(x_c)$ is a definite value. But there are 2^5 possible LBP patterns corresponding to this $\theta(x_c)$. Therefore, LBP retains more local structure information than the Orientation operator. Besides, $\theta(x_c)$ is just segmented into several directions, which is not an effective way to represent local features. Hence, benefitting from the high discriminability of LBP, WLBP is more discriminative than WLD.

Though LBP is highly discriminative, the extension to the LBP operator based on "uniform" patterns still has some shortcomings: it is sensitive to noise and discards some important texture information [35]. In such local patterns selection approach, all non-uniform patterns are simply merged into a single pattern. This simple mergence definitely makes it lose some local features. The basic idea behind the LBP approach is the description of a local pattern via a hard thresholding scheme (Eq. (14)), which makes it sensitive to small gray-level perturbations or noise. In some extreme situations such as dramatic changes in light or noise, the robustness of LBP declines drastically. While WLBP can suppress the influence of noise, as the differential excitation component has smoothing effect by using Laplacian of Gaussian

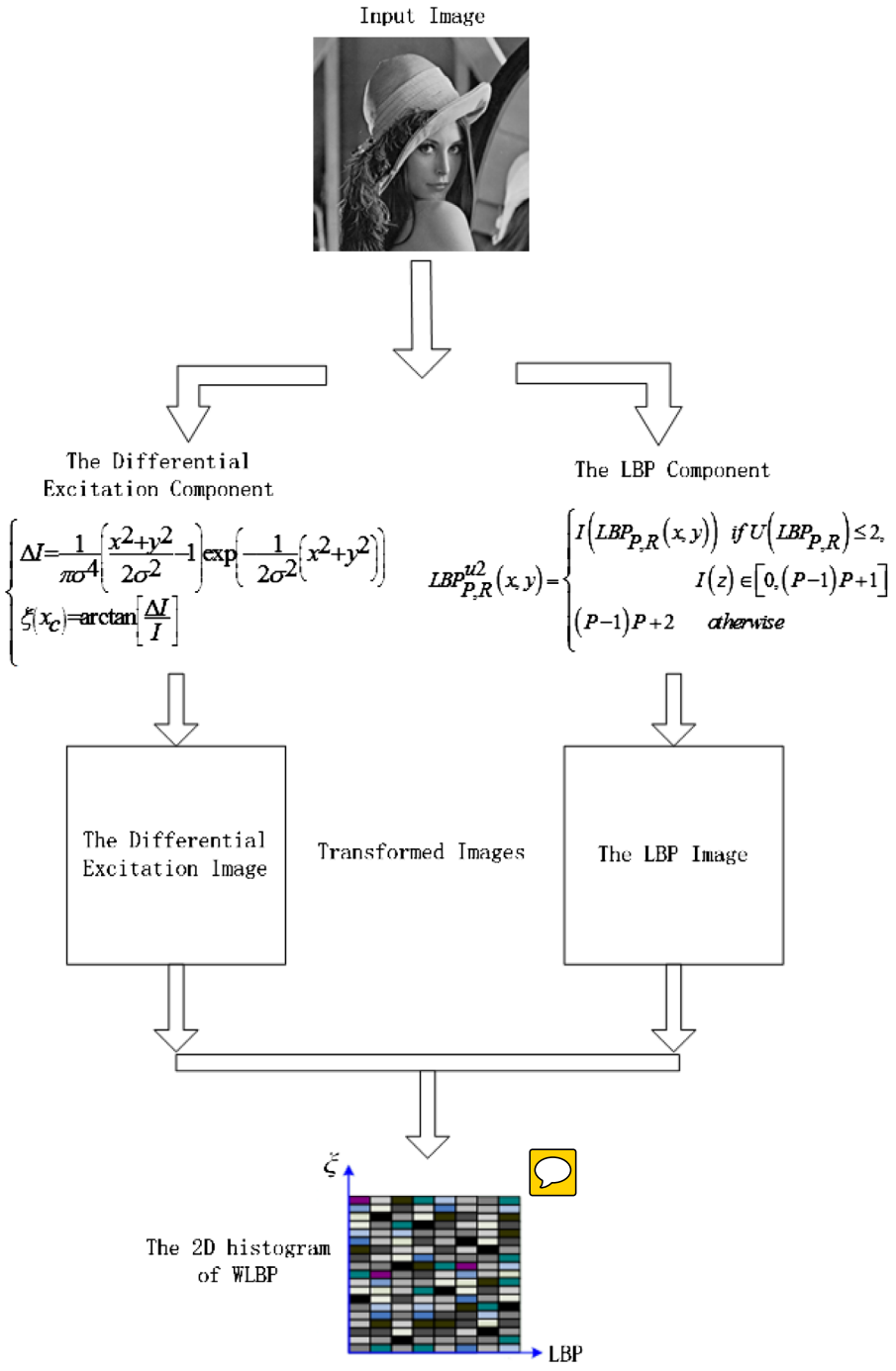


Fig. 2. Illustration of the computation of the WLBP descriptor.

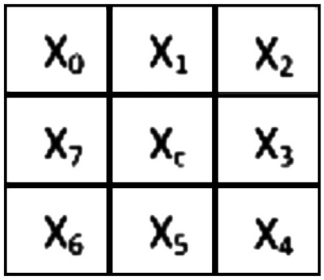


Fig. 3. A pixel with its 8 neighbors.

(LoG). In addition, based on Eq. (9), the sum of p-neighbors' differences to current pixel is divided by the intensity of current pixel, which decreases the influence of multiplicative noise. Meanwhile, a brightness change in which each pixel value is multiplied by a constant will also be canceled by the division.

For WLBP, the differential excitation is divided into high perception interval and low perception interval corresponding to the larger value and smaller value of ξ. Experiments demonstrate that WLBP inherits the high discriminative performance of LBP because the low perception interval contains a majority of "uniform" patterns of LBP. While the larger differential excitation provides a majority of detail information, which enhances the

discriminability further. Hence, theoretically speaking, WLBP is more discriminative than or equal to LBP. By combining the differential excitation and LBP, WLBP not only holds the high discriminability but also obtains the robustness to noise and illumination.

3. Experiments for face recognition

In this section, we use the WLBP histogram feature for face recognition and compare the performance with those of the state-of-the-art methods. The experiments are carried out on two well-known face image databases (FERET and AR).

3.1. Background

Automatic face recognition has become a very active topic in computer vision research for recent two decades [5], which has potential application values as well as theoretical challenges. As a typical pattern recognition problem, the key issue in face recognition is finding efficient features for face representation. Up to now, many facial feature design and extraction methods have been proposed, which can be classified into two categories: global features and local features.

The conventional global feature extraction methods include Principal Component Analysis (PCA) [6], Kernel Principal Component Analysis (KPCA) [7] and Two Dimensional Principal Component Analysis (2DPCA) [8]. PCA is a well-known method for feature extraction. By calculating the eigenvectors of the covariance matrix of the original inputs, PCA linearly transforms a high-dimensional input vector into a low-dimensional one whose components are uncorrelated. Nonlinear PCA has been developed by using different algorithm [9]. As one of them, KPCA is developed by generalizing the kernel method into PCA [10]. 2DPCA is developed for image feature extraction. As opposed to conventional PCA, 2DPCA is based on 2D matrices rather than 1D vectors. That is, the image matrix does not need to be previously transformed into a vector. Instead, an image covariance matrix is constructed directly using the original image matrices, and its eigenvectors are derived for image feature extraction [8]. All the above methods considering the global information of face images could achieve high accuracy under the limited conditions. But the performance degrades with variations in facial expression, illumination condition, pose, partial occlusion, etc., which the global features are sensitive to.

In recent years, local features have gained much attention due to their robustness to challenges such as expression and illumination changes. As one of the most important local features, Local Binary Pattern (LBP) has already been used in various texture and face image analysis tasks [1], which contains abundant texture information about the distribution of some local micro-patterns, such as edges, spots and flats. It was successfully introduced in face recognition field by Ahonen et al. [11]. The face image is divided into several regions from which the LBP histograms are extracted and concatenated into an enhanced feature vector to be used as a face descriptor [11]. WLD was proposed as a simple and powerful local descriptor by Chen et al. [2], which has been applied in face detection and texture classification. In this paper, WLD and our proposed WLBP are applied in face recognition.

3.2. Face description with WLBP

As texture descriptors, LBP, WLD and WLBP consider the small-scale relationships about image pixels and tend to average over the image area. This is desirable property for ordinary textures, because the small-scale relationships determine the appearance the texture. But for faces, the situation is different: retaining the information about spatial relations is important [11]. This reasoning leads us to adopt the idea of dividing facial image into local regions. And local descriptors are extracted from each region independently. As shown in Fig. 4, it is an illustrated image that has been divided into local regions.

As the facial image is divided into $M \times N$ regions, a histogram is computed independently within each region. For the region $R_{i,j}$, $(i = 1, 2, \ldots, M, j = 1, 2, \ldots, N)$, we first compute the 2D histogram $\{WLBP(s,t)\}_{i,j}$ and convert it to 1D histogram $H_{i,j}$, $(i = 1, 2, \ldots, M, j = 1, 2, \ldots, N)$. The $M \times N$ histograms are combined yielding the spatially enhanced histogram. In the spatially enhanced histogram, we have a effective description of the face on three different levels of locality: the LBP labels and the differential excitation values for the histogram contain information about the patterns on a pixel-level, the histogram over a small region produce information on a regional level and the regional histograms are concatenated to build a global description of the face.

To perform the face recognition, there are two essential issues: face representation and classifier design. For a given image, we extract the WLBP histogram to represent the face image. Here, we experimentally set $S = 5$, (8, 2) neighborhood and divide the facial image into 4×4 regions. This is the result of the tradeoff between discriminability and statistical reliability. In a real application, the

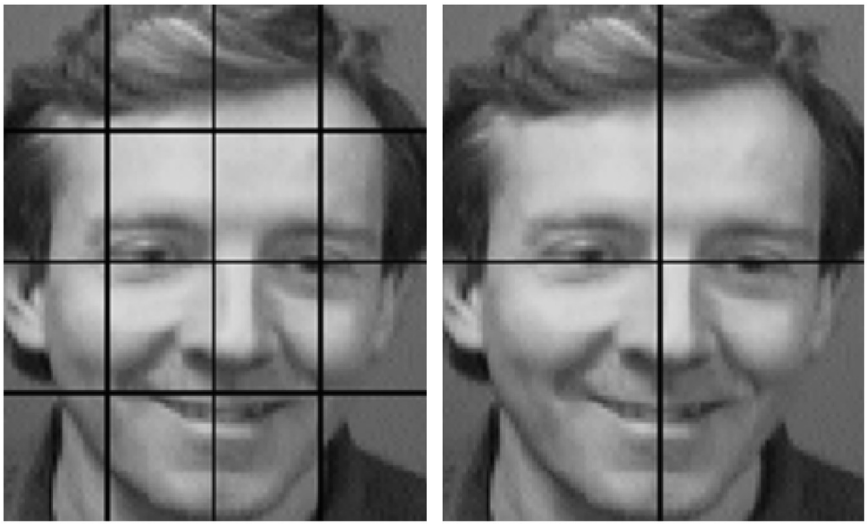


Fig. 4. A facial image divided into 4∗4, 2∗2 rectangular regions.

size of the input image is fixed. If these parameters (i.e., S, the size of neighborhood and the region number) become larger, the dimensionality of the histogram (i.e., the number of its bins) becomes larger and the entries of each bin become smaller. So, the discriminability of the histogram increases while the statistical reliability decreases. In contrast, if these parameters become smaller, the histogram becomes statistically more reliable but the discriminability decreases. To compute the distance of two histograms, the Chi square statistic (χ^2) similarity measurement is used, which is shown in Eq. (15). Finally, the nearest neighbor (NN) classifier is employed for classification.

Chi square statistic (χ^2):

$$\chi^2(H^1,H^2)=\sum_i \frac{(H_i^1-H_i^2)^2}{H_i^1+H_i^2} \tag{15}$$

3.3. Experiments on the FERET database

The FERET face database is the result of FERET program, which was sponsored by the US Department of Defence through the DARPA program [33,34]. It has become a standard database for testing and evaluating state-of-the-art face recognition algorithms. The proposed method was tested on a subset of the FERET database. This subset includes 1400 images of 200 individuals (each individual has seven images). This subset involves variations in facial expression, illumination, and pose. In our experiment, each image in FERET database was manually cropped and resized to 80 × 80. Seven sample images of one person are shown in Fig. 5. In our experiments, we randomly selected four images of each subject for training and the others for testing. All experimental results are average value over ten runs.

3.3.1. Evaluation without noise

As shown in Fig. 6, comparative performance is carried out against some classical face recognition schemes, such as PCA, KPCA and 2DPCA, which extract global features. For LBP, WLD and WLBP, a facial image was divided into 4 × 4windows and (8, 2) neighborhood was selected. The experimental results show that our method outperforms other methods. Although the images in the FERET face database are subject to complex nonlinear changes due to large pose, expression and illumination variations, our proposed method works well in the complex circumstance.

Fig. 5. Seven sample images of one person in FERET face database.

3.3.2. Evaluation with noise

For the comparison of LBP, WLD and WLBP method, further experiment is conducted on the FERET database with added white Gaussian noise. The results obtained at different noise levels for the three methods are illustrated in Table 1. Here, the column specifies the power of additive noise matrix in decibels relative a watt. The noise matrix is added to the original dataset. In addition, the parameter K, S and σ were fixed at pi/30, 5 and 0.6.

It can be noticed that the performance of the three methods degrades as the noise strength increases. However, WLBP outperforms the LBP and WLD for every noise level. When the noise strength reaches a certain level, the advantage of WLBP gets greater. The main reasons lie in the following factors: Gaussian smoothing filter and the computation means of differentia excitation using the sum of p-neighbor differences and the ratio to the current pixel.

3.3.3. The effects of parameters

In this section, we discuss the influence of the parameter selection of K and S and the standard deviation σ in Eq. (8). K is the threshold to segment the differential excitation into high perception interval and low perception interval. $[-K,K]$corresponds to the low perception interval, which contains a majority of "uniform" patterns of LBP. In Table 1, when K ranges from pi/40 to pi/2, the number of

Table 1
The accuracy rate (%) of noise sensitivity evaluation.

	SNR							
	0	5	10	15	20	25	30	35
LBP	82.07	78.3	72.97	60.55	35.12	9.9	2.03	0.9
WLD	87.58	78.8	68.95	53.7	34.05	18.03	10.8	6.07
WLBP	90.12	83.15	74.93	62.05	43.02	28.23	20.22	10.57

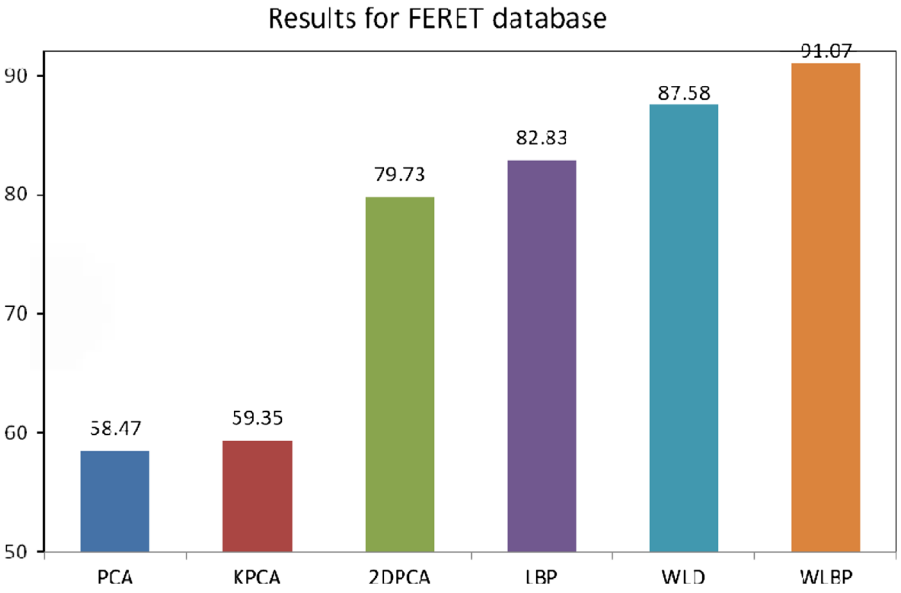
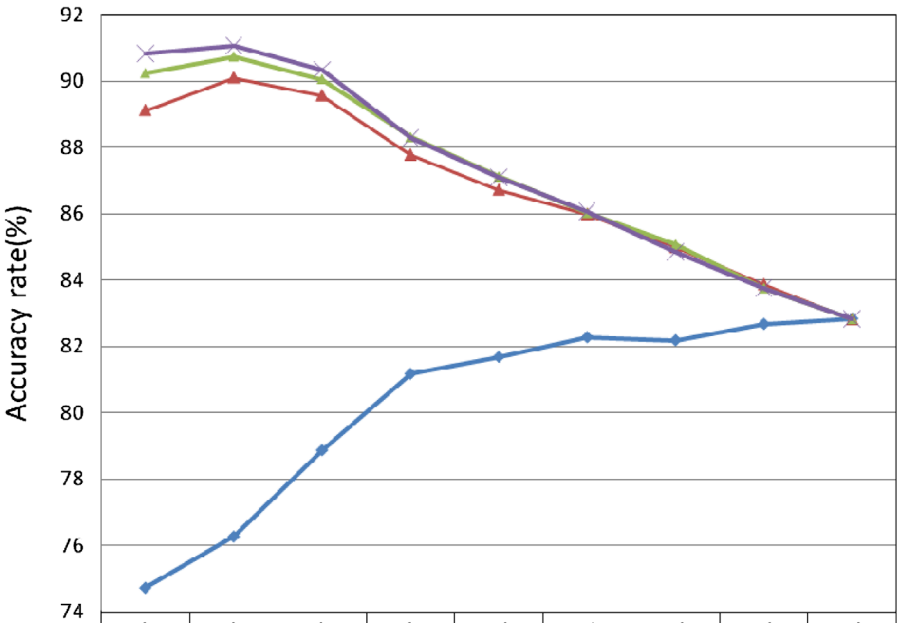


Fig. 6. Results comparison with state-of-the-art methods on FERET database.



	pi/40	pi/30	pi/20	pi/10	pi/8	pi/6	pi/4	pi/3	pi/2
S=1	74.73	76.27	78.87	81.17	81.68	82.27	82.18	82.67	82.83
S=3	89.12	90.1	89.57	87.78	86.73	86	85	83.88	82.83
S=5	90.25	90.75	90.07	88.32	87.13	86.03	85.08	83.77	82.83
S=7	90.83	91.07	90.35	88.28	87.1	86.07	84.85	83.77	82.83

Fig. 7. The effects of using different K and S parameters when $\sigma$is fixed at 0.6.

Table 2
The ratio (%) of "uniform" patterns belonging to the low perception interval $[-K, K]$.

σ	K								
	pi/40	pi/30	pi/20	pi/10	pi/8	pi/6	pi/4	pi/3	pi/2
$\sigma=0.4$	16.4	21.46	30.8	**52.54**	60.27	69.79	81.84	89.89	100
$\sigma=0.5$	34.84	43.62	**57.01**	76.83	81.61	86.88	92.98	96.57	100
$\sigma=0.6$	56.7	**65**	75.51	87.88	90.75	93.81	97.01	98.59	100
$\sigma=0.7$	**68.69**	75.77	83.53	92.56	94.48	96.38	98.25	99.20	100

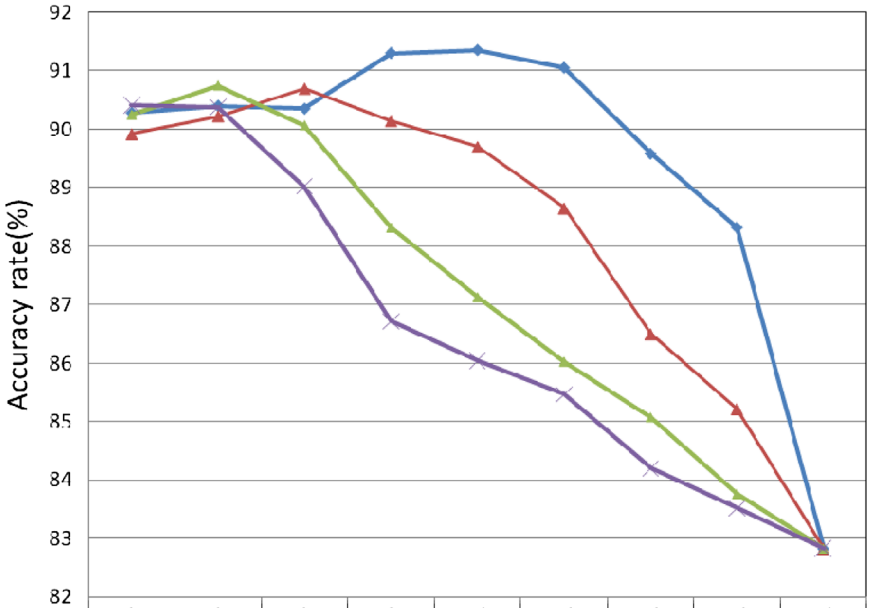
"uniform" patterns belong to this interval gets greater until 100%. Therefore, in Fig. 7, the curve of $S=1$ rises gradually when K increases from pi/40 to pi/2. In our method, the high perception interval is divided further. S is the number of all intervals, which, according to the symmetry, is an odd. When $S>1$, as the high perception intervals contain a large number of detail information, the performance is far better than that of $S=1$, which just considers the low perception patterns. But it degrades as K increases.

Because the high perception intervals become smaller with K increasing and the size of the input image/patch is fixed, the number of pixels belonging to these intervals decreases, which can be seen in Table 2. If the number becomes too small, the statistical reliability of the histogram cannot be guaranteed, which degrades the discriminability. When K becomes small, the discriminability of the low perception interval is limited as the "uniform" patterns contained in this interval drops. This is the reason why the performance declines at $K=$pi/40 in Fig. 7. For a histogram based method, the setting of K is actually a tradeoff between discriminability and statistical reliability. In Fig. 8, the peaks of four curves respectively lie at $K=$pi/10, $K=$pi/20, $K=$pi/30, and $K=$pi/40, which correspond to the bold numbers (52.54%, 57.01%, 65%, 68.69%) in Table 2. Because these settings can balance the discriminability and statistical reliability exactly.

From the curves corresponding to $S>1$ in Fig. 7, one can discover that the performance changes slightly when S varies from 3 to 7. Therefore, to balance dimensionality and performance, we set $S=5$,$K=$pi/30. In Fig. 8, when the value of K is larger than pi/20, the performance degrades as σ increases. The reason is that some details of image will be lost due to the effect of Gaussian smoothing filter. From Figs. 7 and 8, we also note an interesting phenomenon that all the curves meet at the same point at $K=$pi/2. This is because WLBP equals LBP when $K=$pi/2.

3.4. Experiments on the AR database

In this experiment, we use the same dataset and experimental setup as that used in paper [8]. The AR face database [12,13] contains over 4000 color face images of 126 people (70 men and 56 women), including frontal views of faces with different facial expressions, lighting conditions and occlusions. The pictures of most persons were taken in two sessions (separated by two weeks). Each section contains 13 color images and 120 individuals (65 men and 55 women) participated in both sessions. The images of these 120 individuals were selected and used in our experiments. Only the full facial images were considered here (no attempt was made to handle occluded face recognition in each session). We manually cropped the face portion of the image and then normalized it to 50×40 pixels. The normalized images of one person are shown in Fig. 9, where image (1)–(7) are from Session 1, and images (14)–(20) are from Session 2. The details of the images are (1), neutral expression; (2), smile; (3), anger; (4), scream; (5), left light on; (7), right light on; (7), all sides light on; and images (14)–(20) were taken under the same conditions as (1)–(7).



	pi/40	pi/30	pi/20	pi/10	pi/8	pi/6	pi/4	pi/3	pi/2
σ=0.4	90.28	90.4	90.35	91.3	91.35	91.05	89.58	88.32	82.83
σ=0.5	89.92	90.23	90.7	90.15	89.7	88.65	86.5	85.22	82.83
σ=0.6	90.25	90.75	90.07	88.32	87.13	86.03	85.08	83.77	82.83
σ=0.7	90.42	90.38	89.03	86.72	86.05	85.47	84.2	83.52	82.83

Fig. 8. The effects of using different K and σ parameters when S is fixed at 5.

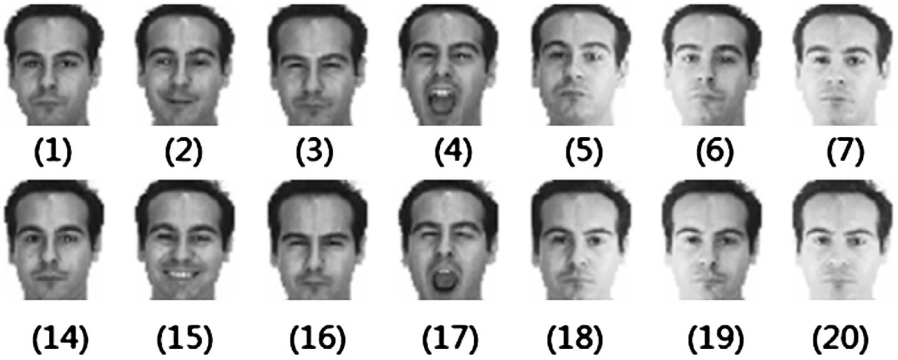


Fig. 9. Samples images for one subjects of the AR database.

3.4.1. Variation over time

In this experiment, images from the first session (i.e., (1)–(7)) were used for training, and images from the second session (i.e., (14)–(20)) were used for testing. Thus, the total number of training samples was 840. Since the two sessions were separated by an interval of two weeks, the aim of this experiment was to compare the performance of WLBP and the other methods (PCA, KPCA, 2DPCA, LBP, WLD) under the conditions where there are changes over time.

3.4.2. Variation in facial expressions

In this experiment, the objective was to compare WLBP and the other methods (PCA, KPCA, 2DPCA, LBP, WLD) under varying facial expressions. We selected images (1)–(4) and (14)–(17), which involve variations in facial expressions. Images (1) and (14) were used for training and the others (i.e., (2)–(4) and (15)–(17)) were used for testing. Thus, the total number of training samples is 240.

3.4.3. Variation in lighting conditions

In this experiment, our aim was to test the robustness of WLBP under varying illumination. Images with varying lighting conditions were selected first. The selected samples set included images (1), (5)–(7) from the first session and images (14), (18)–(20) from the second session. From this set, we arbitrarily chose two samples for training, one from the first session and another from the second. The remaining samples were used for testing. Thus, there were 16 possible sets of training samples. Based on these sets of training samples, we performed 16 tests and got the average.

3.4.4. Analysis of the experimental results

The results of the three experiments mentioned above are shown in Fig. 10, which show that WLBP works in a very robust way under different challenges (time, lighting and expression). Although recognizing duplicate faces which are taken two weeks later is a challenge, our proposed method still shows its impressive superiority. Under the change of time, our method achieved the

best performance with a recognition rate of 92.98% against 66.55%, 67.14%, 67.74%, 88.57% and 87.38% for PCA, KPCA, 2DPCA, LBP and WLD. Though all the mentioned methods are robust to the variation of facial expression, the performance of WLBP is the best. Additionally, our method is robust to variation of illumination which obtained the best recognition rate of 96% under different lightings, while the best performance of 2DPCA was just 84.52%. WLBP has been developed to suppress the effects of illumination change. For LBP component, it compares its neighbor pixels with current pixel. So, a brightness change in which a constant is added to each image pixel will not affect the pattern values. On the other hand, for the differential excitation component, it performs the division between the differences ΔI and the current pixelx_c. Thus, a change in image contrast in which each pixel value is multiplied by a constant will be canceled by the division.

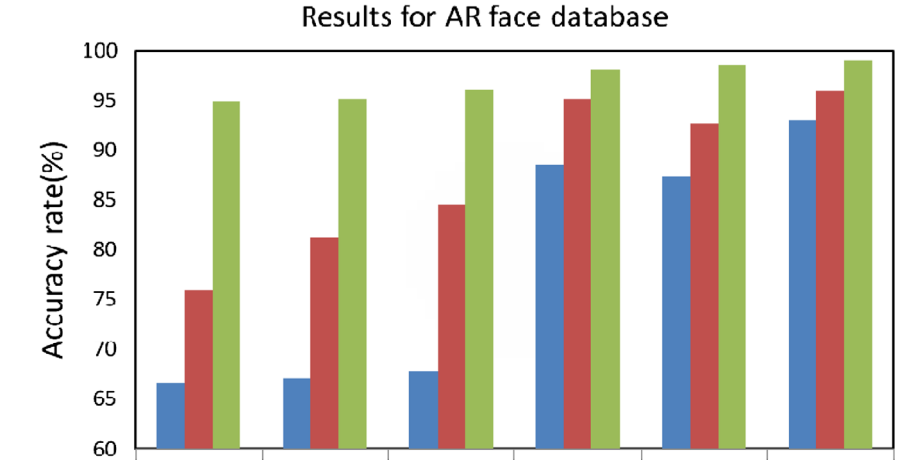
4. Experiments for texture classification

Texture classification is a fundamental problem in computer vision with a wide variety of applications [14]. In this section, we use the WLBP histogram feature for texture classification and two experiments were performed to test our approach. The experiments are respectively carried out on the Brodatz [15] and KTH-TIPS2-a [16] texture databases.

4.1. Datasets and experimental setup

Examples of the 32 Brodatz [17] textures used in the experiments are shown in Fig. 11(a). The images are 256×256 pixels in size and have 256 Gy levels. Each image was divided into 16 disjoint samples of size 64×64 pixels, which were independently histogram-equalized to remove luminance differences between textures. To make the classification problem more challenging and generic, three additional samples were generated from each sample: (i) a sample rotated by 90°, (ii) a 64×64 scaled sample obtained from the 45×45 pixels in the middle of the original sample, and (iii) a sample that was both rotated and scaled. Consequently, the classification problem involved a total of 2048 samples, 64 samples in each of the 32 texture categories. The KTH-TIPS2-a database contains 4 physical, planar samples of each of 11 materials under varying illumination, pose and scale. The database contains images at nine scales, under four different illumination directions, and three different poses. There are 11 texture classes with 4395 images in all. The images are 200×200 pixels in size (we did not include those images which are not of this size), and they are transformed into 256 Gy levels. Some examples from each sample are shown in Fig. 11(b).

For the two texture experiments, we use the same setup as that used in the paper [2]. For Brodatz textures, experiments are carried out with ten-fold cross validation to avoid bias. For each round, we randomly divided the samples in each class into two subsets of the



	PCA	KPCA	2DPCA	LBP	WLD	WLBP
Time	66.55	67.14	67.74	88.57	87.38	92.98
Lighting	75.87	81.27	84.52	95.07	92.63	96
Experssion	94.86	95.14	96.11	98.06	98.61	99.03

Fig. 10. Comparison of WLBP with other methods using the AR database.

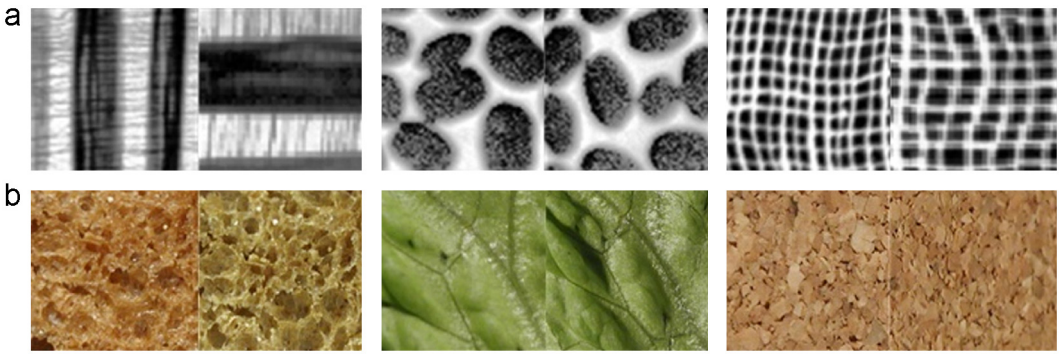


Fig. 11. Some examples from two texture datasets (a) Brodatz and (b) KTH-TIPS2-a.

same size, one for training and the other for testing. In this fashion, the images belonging to the training set and to the test set are disjoint. The result is the average value over the ten runs. For KTH-TIPS2-a textures, all images of three samples are for training, while all images of the remaining sample are for testing. All images for training or testing are at one certain scale. This experiment is repeated four times by randomly selecting different three samples for training. The result is reported as the average value over the four runs.

4.2. Experimental results

As shown in Fig. 12, we compare our method with others on the classification task of Brodatz and KTP-TIPS2-a texture: SIFT, LBP, WLD, MLBP and WLBP. Note that the result of SIFT is quoted directly from the paper [2]. The result of WLD is not the same as that of the paper [2], because the nearest neighbor (NN) classifier was selected instead of K-nearest neighbor classifier. And we use the Chi square statistic (χ^2) similarity measurement instead of the normalized histogram intersection. In addition, to simplify the calculation of WLD, we set $M=6$, $T=8$, $S=1$, instead of $M=6$, $T=8$, $S=20$ in the paper [2]. MLBP denotes $LBP_{8,1}+LBP_{16,2}$. The feature of MLBP for each image is to concatenate the histograms from different operators which are realized with different (P,R) as discussed in Section 2. Although the samples from Brodatz textures are rotated and scaled, we still obtained satisfactory results. It shows that WLBP is robust to rotation and scaling. For the KTH-TIPS2-a dataset, the variations are much more diverse than those of the Brodatz set. Therefore, the accuracies of all the tested methods (SIFT,LBP,WLD,MLBP) are lower than those on Brodatz. But the performance of WLBP is still favorable. As shown in Fig. 11(b), it is superior to all the other descriptors in this paper.

5. Conclusion

In this paper, we propose a local descriptor called WLBP, which consists of two components: differential excitation and LBP. The differential excitation component is derived from Weber's law, which can extract salient micro-patterns by simulating the perception of human beings. We improve the different excitation operator by bringing in Laplacian of Gaussian. Due to the high discriminability of LBP, we adopt the LBP operator instead of the Orientation operator in WLD. By encoding patterns of differential

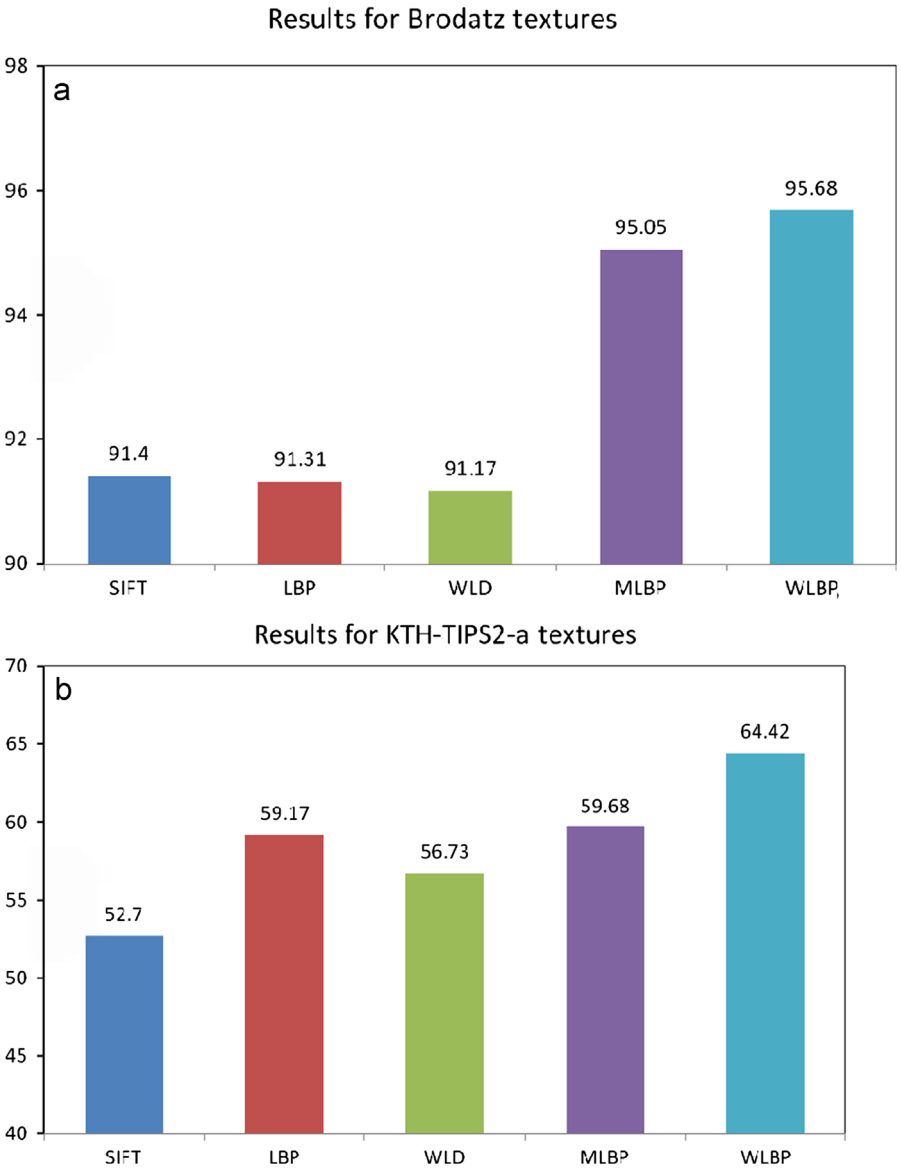


Fig. 12. Results comparison on Brodatz and KTH-TIPS2-a textures.

excitation and LBP, we compute the 2D histogram of WLBP, which is more discriminative than LBP and WLD.

The proposed descriptor is accessed with the face recognition task. Experimental results show that WLBP is superior to those state-of-the-art global methods (e.g., PCA, KPCA, 2DPCA). The performance of WLBP is better than LBP and WLD not only in noise free images but also in the images with noise. In addition, WLBP is robust to different challenges (time, facial expression and lighting). Experiments on texture databases also demonstrate that WLBP is a powerful texture descriptor, which outperforms SIFT, LBP, WLD and MLBP.

Although the performance of WLBP has already been excellent, some improvements are still possible. Indeed, the feature vector length of WLBP is obviously longer than that of LBP and WLD. It slows down the recognition speed, especially for very large databases. In our future work, we will consider to reduce dimensionality of the feature vector and improve the computing efficiency. In addition, future interest also lies in how to exploit the proposed descriptor for the application of object recognition.

References

[1] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray scale and rotation invariant texture analysis with local binary patterns, IEEE Trans. Pattern Anal. Mach. Intell. 24 (7) (2002) 971–987.

[2] Jie Chen, Shiguang Shan, Chu He, et al., WLD: a robust local image descriptor, IEEE Trans. Pattern Anal. Mach. Intell. 32 (9) (2010) 1705–1720.

[3] A.K. Jain, Fundamentals of Digital Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1989.

[4] JiChao Jiao, Baojun Zhao, A new feature descriptor and selection method to space image registration, in: IASP, International Conference, 2010, pp. 314–317.

[5] S.Z. Li, A.K. Jain (Eds.), Springer, 2005.

[6] M. Turk, A. Pentland, Eigenfaces for recognition, J. Cognitive Neurosci. 3 (1) (1991) 71–86.

[7] K. Etemad, R. Chellappa, Discriminant analysis for recognition of human face images, J. Opt. Soc. Am. 14 (1997) 1724–1733.

[8] J. Yang, D. Zhang, A.F. Frangi, J. Yang, Two-dimensional PCA:A new approach to appearance-based face representation and recognition, IEEE Trans. Pattern Anal. Mach. Intell. 26 (1) (2004) 131–137.

[9] K.I. Diamantaras, S.Y. Kung, Principal Component Neural Networks, Wiley, New York, 1996.

[10] B. Scholkopf, A. Smola, K.R. Muller, Nonlinear component analysis as a Kernel eigenvalue problem, Neural Comput. 10 (1998) 1299–1319.

[11] Matti Pietikäinen Timo Ahonen, Face description with local binary patterns: application to face recognition, IEEE Trans. Pattern Anal. Mach. Intell. 28 (12) (2006) 2037–2041.

[12] A.M. Martinez, R. Benavente, The AR Face Database, CVC Technical Report, 24, 1998.

[13] A. M. Martinez, R. Benavente, The AR Face Database ⟨http://rv.l1.ecn.purdue.edu/~aleix/aleix_face_DB.html⟩, 2003.

[14] Texture analysis, in: ,C.H. Chen, L.F. Pau, P.S.P. Wang (Eds.), Handbook of Pattern Recognition and Computer Vision, World Scientific, Singapore, 1993, pp. 235–276.

[15] P. Brodatz., Textures: A Photographic Album for Artists and Designers, Dover Publications, New York, 1966.

[16] B. Caputo, E. Hayman and P. Mallikarjuna, Class-Specific Material Categorisation, in: Proceedings of the IEEE International Conference on Computer Vision, 2005.

[17] T. Ojala, K. Valkealahti, E. Oja, M. Pietikäinen, Texture discrimination with multidimensional distributions of signed gray level differences, Pattern Recognition 34 (3) (2001) 727–739.

[18] V. Bruni, D. Vitulano, A Generalized Model for Scratch Detection, IEEE Trans. Image Process. 13 (1) (2004) 44–50.

[19] T. Phiasai, P. Kumhom K. Chamnongthai, A digital image watermarking technique using prediction method and Weber ratio, in: International Symposium on Communications and Information Technologies, 2004.

[20] F. Schaffalitzky, A. Zisserman, Multi-view matching for unordered image sets, in: Proceedings of the Seventh European Conference on Computer Vision, 2002, pp. 414–431.

[21] Wei Li Yanwei Pang, Yuan Yuan, Jing Pan, Fully affine invariant SURF for image matching, Neurocomputing 85 (2012) 6–10.

[22] Jing Li, Nigel M. Allinson, A comprehensive review of current local features for computer vision, Neurocomputing 71 (2008) 1771–1787.

[23] P. Moreels, P. Porena, Evaluation of features detectors and descriptors based on 3D objects, Int. J. Comput. Vis. 73 (3) (2007) 263–284.

[24] K. Mikolajczyk, C. Schmid, A performance evaluation of local descriptors, IEEE Trans. Pattern Anal. Mach. Intell. 27 (10) (2005) 1615–1630.

[25] David G. Lowe, Distinctive image features from scale invariant keypoints, Int. J. Computer Vis. 60 (2) (2004) 91–110.

[26] N. Dalal, B. Triggs, Histograms of oriented gradients of human detection, in: CVPR, 2005.

[27] Y. Pang, Y. Yuan, X. Li, J. Pan, Efficient HOG human detection, Signal Process. 91 (4(April)) (2011) 773–781.

[28] Y. Pang, H. Yan, Y Yuan, K. Wang, Robust CoHOG Feature extraction in human-centered image/video management system, IEEE Trans. Syst. Man Cybern. Part B: Cybern. 42 (2) (2012) 458–468.

[29] 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, 2005, pp. I:786–791.

[30] J. Luo, Y. Ma, E. Takikawa, S. Lao, M. Kawade, B. Lu, Person-specific SIFT features for face recognition, in: Proceedings of the ICASSP, vol. 2, 2007, pp. 593–596.

[31] S. Liao, M.W.K. Law, C. Schmid, Dominant local binary patterns for texture classification, IEEE Trans. Image Process. 18 (no. t) (2009) 1107–1118.

[32] B. Manjunath, W. Ma, Texture features for browsing and retrieval of image data, IEEE Trans. Pattern Anal. Mach. Intell. 18 (8) (1996) 837–842.

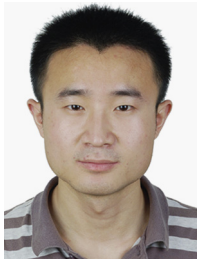
[33] P.J. Phillips, H. Moon, S.A. Rizvi, P.J. Rauss, The FERET evaluation methodology for face-recognition algorithms, IEEE Trans. Pattern Anal. Mach. Intell. 22 (10) (2000) 1090–1104.

[34] P.J. Phillips, The Facial Recognition Technology (FERET) Database, 2004 ⟨http://www.itl.nist.gov/iad/humanid/feret/feret_master.html⟩.

[35] Runsheng Hui Zhou, Wang, Cheng Wang, A novel extended local binary pattern operator for texture analysis, Inf. Sci. 178 (22) (2008) 4314–4325.

Jinhui Tang is currently a Professor of School of Computer Science and Technology, Nanjing University of Science and Technology. He received his B.E. and Ph.D. degrees in July 2003 and July 2008 respectively, both from the University of Science and Technology of China (USTC). From July 2008 to December 2010, he worked as a research fellow in School of Computing, National University of Singapore. During that period, he visited School of Information and Computer Science, UC Irvine, as a visiting research scientist, from January 2010 to April 2010. His current research interests include large-scale multimedia search, social media analysis and computer vision. He has over 60 journal and conference papers in these areas.

Dr. Tang served as a guest editor of IEEE Transactions on Multimedia, ACM Transactions on Intelligent Systems and Technology, ACM/Springer Multimedia Systems Journal, and Springer Journal of Multimedia Tools and Applications, a co-chair of the ACM the First International Workshop on Web-Scale Multimedia Corpus 2009, a Technical Committee Member for about 30 international conferences, such as SIGIR and ICME, and a reviewer for about 30 prestigious international journals. Dr. Tang is a recipient of the 2008 President Scholarship of Chinese Academy of Science, and a co-recipient of the Best Paper Award in ACM Multimedia 2007. He is a member of ACM, IEEE and IEEE-CS.



Fan Liu is currently a Ph.D. candidate of School of Computer Engineering of Nanjing University of Science and Technology. He received his B.S. degree in network engineering from Nanjing University of Science and Technology (NUST) in June 2005. He is currently working towards the Ph.D. degree in computer application technology at Nanjing University of Science and Technology. His research interests include computer vision, image processing and pattern recognition.

Zhenmin Tang is currently a professor of School of Computer Engineering, Nanjing University of Science and Technology. He received his B.S. degree from Harbin Engineering University (HEU) in January 1982. He received his M.S. and Ph.D. degrees in January 1988 and March 2003 respectively, both from Nanjing University of Science and Technology (NUST). He is currently the president of School of Computer Engineering, Nanjing University of Science and Technology. His current research interests are in the area of image processing, pattern recognition, artificial intelligence and expert system. He has over 60 journal and conference papers in these areas.