

Fusing Facial Texture Features for Face Recognition

Yanqing Shao^{1,2} · Chaowei Tang¹ · Min Xiao³ · Hui Tang⁴

Received: 3 January 2014 / Revised: 28 December 2015 / Accepted: 18 April 2016 / Published online: 9 May 2016
© The National Academy of Sciences, India 2016

Abstract Aiming at taking full advantage of facial information both in low-frequency and high-frequency regions and further improving face recognition rate, this paper constructs a robust nonsubsampling contourlet transform local binary patterns (NSCTLBP) feature and proposes a face recognition method fusing NSCTLBP and Gabor features. Firstly, face image is decomposed by NSCT, and the LBP values of NSCT high-frequency subbands are computed to construct NSCTLBP features. Meanwhile, convolution of 2D-Gabor wavelet with face image is performed to extract Gabor texture feature in low-frequency. Secondly, Euclidean distance and eigenvalue-weighted cosine (EWC) distance are adopted to explore the similarity measurement of NSCTLBP and Gabor features respectively. Finally, the face images are matched according to the weighted similarity of NSCTLBP feature and Gabor feature collaboratively. Experimental results on Yale and ORL databases show that the proposed method has better performances than that based on NSCT feature, NSCTLBP feature and Gabor feature separately against illumination, expression, and angle variations and glasses occlusion.

Keywords Eigenvalue-weighted cosine distance · Face recognition · Gabor wavelet · Local binary patterns · Nonsubsampling contourlet transform

1 Introduction

Face recognition technology has been widely used in many fields. However, in reality, the captured face images may suffer from variations due to uncontrolled acquisition such as illumination, expression, angle variations and glasses occlusion, which affects the efficiency of face recognition system.

Researchers have spared no efforts to obtain some robust and effective face recognition methods. Texture feature based face recognition methods have attracted more attention. Texture information of face image can be utilized to measure intensity distribution (i.e. gray values of gray image) in different situations (i.e. wrinkles, bumps and dents). Eleyan and Demire used grey level co-occurrence matrix (GLCM) operator to describe image texture information about intensity distribution and relative position of neighborhood pixels [1]. Yu et al. [2] enhanced the method by combining GLCM and weighted Euclidean distance. Local binary pattern (LBP) operator [3] has been widely used to extract texture details of face image [4, 5], but when the resolution of image changes, the calculated textures are not accurate. Zhang et al. [6] proposed high-order local derivative patterns (LDP) descriptor by encoding distinctive spatial relationship in given regions rather than the relationship between the central point and its neighbors in LBP. This method is claimed to be more effective than LBP, but the size of extracted features is large. Meththa et al. [7] extracted directional and textural feature by

✉ Chaowei Tang
cwtang@cqu.edu.cn

¹ College of Communication Engineering, Chongqing University, Chongqing 400044, China

² Communication Engineering Department, Chongqing College of Electronic Engineering, Chongqing 401331, China

³ Software and Service Center, Sichuan Changhong Electronic Co., Ltd., Sichuan 621000, China

⁴ High Performance Network Lab, Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190, People's Republic of China

applying modified LBP operator to optimized directional faces. LBP mostly focuses on describing texture details, which leads to poor characterization of shape over a broader range of scale. Wavelet transform was widely used to extract texture information in local area for its property of spatial-frequency localization, multi-scale and multi-orientation [8]. Gabor function which meets the uncertainty principle limit of time–frequency domain is considered as wavelet basis function in 2D-wavelet transform, so it is easier for 2D-Gabor wavelet transform to achieve the best resolution in time–frequency domain [9]. 2D-Gabor with properties of multi-scale and multi-orientation had been widely used for face recognition [10–23]. Since both magnitude and phase of Gabor feature contain rich information, some related approaches and fusion methods have been proposed [11, 13, 14, 16, 17, 20, 22, 24]. Yu et al. [19] had combined Gabor magnitude-based and Gabor phase-based texture representations to construct Gabor magnitude-based and phase-based texture representation for high utilization of Gabor feature. Xia et al. [16] applied block Gabor directed derivative layer local radii-changed binary patterns (BG2D2LRP) to capture static texture differences and dynamic contour trends, more information insensitive to expression interferences can be extracted from BG2D2LRP feature. However, Gabor feature performs poorly in describing appearance details in high-frequency regions.

It is worth noting that most of the approaches, such as local Gabor binary patterns [13], learned local Gabor patterns [21] and histogram of Gabor phase patterns [14, 23], combine Gabor wavelet with other texture descriptors (i.e. LBP) to improve efficiency [13, 14, 16, 18, 20–23]. Gabor feature and LBP feature were fused to gain better performance than either alone [18]. The methods mentioned above can be divided into two categories, one is to fuse Gabor feature and local patterns feature and the other is to utilize local patterns to encode Gabor filtered face image. The former captures most facial information to a certain extent while the latter improves representation capability of local patterns at multiple scales and orientations.

NSCT has the characteristics of shift-invariant, multi-scale, multi-orientation and better directional frequency localization [25]. Therefore, it has favorable performance in capturing contour structure information of signal. Compared with wavelet transform, the support interval of NSCT basis function has an elongated structure which can change the aspect ratio when scale changes. NSCT is anisotropic and it can be adopted to represent texture information with less coefficients. Xu et al. [26] applied NSCT to capture facial contour information of face image and used support vector machine (SVM) to learn and classify NSCT features. The effectiveness was verified by Wang [27]. Xie et al. [28] introduced logarithm transform

into NSCT and then proposed logarithm nonsubsampling contourlet transform. The approach firstly performed logarithm transform on face image, then applied NSCT decomposition on logarithm transformed face image to obtain the low-frequency and high-frequency subbands. Finally illumination invariance was obtained by applying inverse NSCT on the high-frequency subbands processed with Bayes shrink. The extraction of illumination invariant used by Cheng et al. [29, 30] is similar to that used by [28]. The difference is that Xie et al. [28] processes high-frequency subbands with adaptive normal shrink to obtain illumination invariant. However, both the approaches mentioned above only consider the high-frequency subbands, whereas the low-frequency subbands of face image still contain useful facial information. Fan et al. [31] applied histogram equalization to the low-frequency subbands, then performed inverse NSCT on the processed high-frequency subbands and modified low-frequency subbands to obtain more facial information from the extracted illumination invariant. But the multi-orientation analysis of the low-frequency domain is not completed by histogram equalization.

Even though NSCT feature extracted in [26, 28–31] is multi-scale, multi-orientation and shift-invariant, it is not rotation-invariant. The rotation of textures in face image will change the coefficient distribution of subbands of NSCT and reduce recognition rate. To improve the stability of NSCT subband coefficients, a novel feature named NSCTLBP which is invariant to rotation is constructed in this paper. The recognition method in [28–30] utilized high-frequency subbands of NSCT to depict facial textures. Whereas there is still some benefit facial information in low-frequency regions. Studies show that Gabor transform can extract almost all the information in low-frequency regions and NSCT takes the advantage of tree structured filter to perform frequency decomposition of image and extracts rich texture information in high-frequency regions properly. Thus we use NSCTLBP feature to enhance the performance of the method in case of interferences. Moreover, to further improve the utilization of facial information, a weighted measure rule combining Euclidean distance and EWC distance is proposed to fuse NSCTLBP and Gabor features. The proposed method integrates the good high-frequency characterization of NSCTLBP feature and favorable low-frequency performance of Gabor wavelet transform to increase recognition rate. Experimental results on Yale and ORL databases show that the proposed method has better performances than that based on NSCT feature, NSCTLBP feature and Gabor feature alone in case of illumination, expression and angle variations and glasses occlusion.

In later sections we will firstly describe the construction of NSCTLBP feature, and the extraction of the Gabor

feature. Then we will introduce the proposed face recognition method fusing NSCTLBP feature and Gabor feature. Finally, the performance of the proposed method is verified by experiments.

2 Feature Construction and Extraction

2.1 Construction and Extraction of NSCTLBP Feature

Contourlet transform(CT) [32] is an image representation scheme with multi-scale and multi-orientation. It is composed of Laplacian pyramid (LP) and directional filter bank (DFB). CT is not shift-invariant and suffers frequency aliasing due to downsamplers and upsamplers presented in both LP and DFB. NSCT is the shift-invariant version of CT. It consists of nonsubsampling pyramid (NSP) and nonsubsampling directional filter bank (NSDFB). Shift-invariant is adopted to overcome a wide range of changes on subband coefficient distribution caused by image shift. NSP filter ensures NSCT with multi-resolution analysis ability to effectively extract facial information at different scales. While NSDFB guarantees the multi-orientation of NSCT. NSCT decomposition is shown in Fig. 1. Figure 1a displays an overview of NSCT, where NSP filter performs scale decomposition on the original image signals to obtain lowpass image signal and differential bandpass image signals. NSDFB performs tree-structured directional decomposition on the obtained bandpass image signals to get directional subbands. The lowpass image signal obtained previously is continually decomposed into lowpass image signal and differential bandpass image signal of next scale. Repeating the decomposition above, the multi-scale and multi-orientation decomposition of original image is implemented. NSDFB in Fig. 1a splitting bandpass subband into directional subbands is illustrated in Fig. 1b.

NSCT is adopted to decompose the face image with size $M \times N$ into one low-frequency subband coefficient C_0 and high-frequency subband coefficient set $\{C_{1,1}, C_{1,2}, \dots, C_{u,1}, \dots, C_{u,v}\}$ with the same size as face image. $C_{u,v}$ describes facial information in the u -th orientation subband of the v -th scale. LBP calculation is introduced to advance the representation capacity of $C_{u,v}$ and construct NSCTLBP feature with properties of multi-scale, multi-orientation, shift-invariant, rotation-invariant. According to LBP prototype, NSCTLBP feature calculation can be given by [3]

$$C_{u,v-LBP}(x_c, y_c) = \sum_{k=0}^7 s(g_k - g_c)2^{k-1} \tag{1}$$

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

where LBP value of $C_{u,v}$ at point (x_c, y_c) is obtained through binary coding of the coefficients in a neighborhood of 3×3 pixels with center pixel at point (x_c, y_c) . g_c is the value of center coefficient $C_{u,v}(x_c, y_c)$ and g_k is the value of k -th neighborhood coefficient of $C_{u,v}(x_c, y_c)$ in clockwise. LBP value (i.e. $C_{u,v-LBP}^{ri}$) can be calculated as [3]

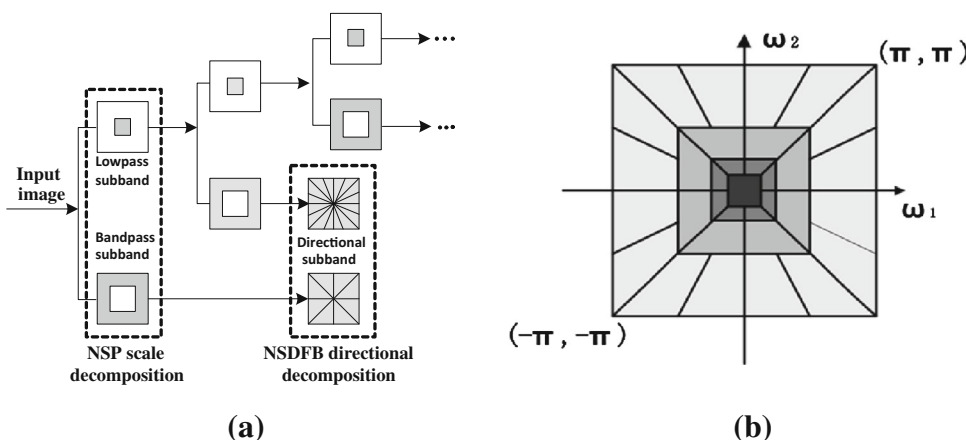
$$C_{u,v-LBP}^{ri} = \min\{ROR(C_{u,v-LBP}, k) | k = 0, 1, \dots, 7\} \tag{2}$$

where $ROR(C_{u,v-LBP}, k)$ performs a circular bit-wise right shift on the 8-bit number $C_{u,v-LBP}$ k times to the right. Finally, the NSCTLBP feature extracted from face image is computed as

$$f_{nscctlbp} = [C_{1,1-LBP}^{ri}, C_{1,2-LBP}^{ri}, \dots, C_{u,1-LBP}^{ri}, \dots, C_{u,v-LBP}^{ri}] \tag{3}$$

As the value at point (x_c, y_c) in each feature vector $C_{u,v-LBP}$ is a minimal circular coding sum of the coefficients in the neighborhood of (x_c, y_c) , the value is stable with the rotation of local textures in face image. Thus NSCTLBP is robust to image rotation. $f_{nscctlbp}$ obtained on the basis of high-

Fig. 1 Nonsubsampling contourlet transform. **a** Nonsubsampling filter bank structure that implements NSCT. **b** Idealized frequency partitioning obtained by NSCT



frequency subband coefficients does not contain low-frequency facial information.

2.2 Extraction of Gabor Wavelet Feature

2D Gabor wavelet is often applied to capture local structure information corresponding to spatial localization, spatial frequency selectivity and orientation selectivity [10]. It enables 2D Gabor wavelet to be widely used for facial textures extraction. 2D Gabor wavelet kernel function is defined by [33]

$$\Psi_{u,v}(z) = \frac{\|k_{u,v}\|^2}{\sigma^2} \exp\left(-\frac{\|k_{u,v}\|^2 \|z\|^2}{2\sigma^2}\right) \left[\exp(ik_{u,v}z) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \tag{4}$$

$$k_{u,v} = \begin{pmatrix} k_v \cos \phi_u \\ k_v \sin \phi_u \end{pmatrix}, \quad k_v = \frac{k_{max}}{f^v}, \quad \phi_u = \frac{\pi u}{8}$$

where $\frac{\|k_{u,v}\|^2}{\sigma^2}$ is used to compensate for energy spectrum attenuation determined by frequency in natural images.

Gaussian envelope function $\exp\left(-\frac{\|k_{u,v}\|^2 \|z\|^2}{2\sigma^2}\right)$ is a window function to make $\Psi_{u,v}(z)$ be locally valid. $\exp(ik_{u,v}z)$ is a complex plane wave, where real part is cosine plane wave while imaginary part is sine plane wave. $\exp\left(-\frac{\sigma^2}{2}\right)$ is adopted to eliminate the effect of DC component of image to Gabor wavelet and ensures $\Psi_{u,v}(z)$ to be insensitive to illumination variations. $z = (x, y)$ denotes the position of pixel, $k_{u,v}$ signifies the filter center frequency, k_v and ϕ_u describe the multi-scale and multi-orientation capability of Gabor filter respectively. By choosing different scales v and orientations u , and appropriate value of k_{max}, σ, f , we will obtain a series of Gabor kernels. From the equation of $k_{u,v}$, the coverage of frequency distribution of 2D Gabor filter is a circular area with radius of k_v . Although 2D Gabor filter is proper in covering the horizontal and vertical high-frequency regions, it performs weak coverage on diagonal high-frequency region of spatial frequency regions of face image.

Gabor feature representation of an image $I(x, y)$ is derived by convolving the image with the 2D Gabor wavelet kernel function at different scales and orientations [34].

$$W_{u,v}(x, y) = I(x, y) * \Psi_{u,v}(x, y) \tag{5}$$

where $W_{u,v}(x, y)$ denotes the output of the convolution of the image $I(x, y)$ with $\Psi_{u,v}(x, y)$ at scale v and orientation u . The adoption of FFT and IFFT to Eq. (5) can speed up the calculation.

$$W_{u,v}(x, y) = \mathcal{F}^{-1}(\mathcal{F}(I(x, y)) \times \mathcal{F}(\Psi_{u,v}(x, y))) \tag{6}$$

where $W_{u,v}$ contains the response of real part and imaginary part of Gabor kernel. The amplitude of $W_{u,v}$ contains local energy variations of image, thus it can be

calculated as image feature. Gabor feature of $I(x, y)$ is generated by combining all the calculated features $W_{u,v}$:

$$f_{Gabor} = (W_{1,1}, W_{1,2}, \dots, W_{u,1}, \dots, W_{u,v}) \tag{7}$$

3 Fusion of NSCTLBP Feature and Gabor Feature

Gabor features and the NSCT features are observed to be complementary to each other to some extent, since Gabor feature mainly extracted from the low frequency region while NSCT feature extracted from the high frequency region of face image. LBP is a fine scale descriptor that captures small texture details and resistant to illumination variations. The combination of NSCT and LBP will be a good choice for coding fine details of facial appearance and texture. Thereby, the complementary nature of Gabor and NSCTLBP makes them be good candidates for fusion. The fusion method can not only achieve complete characterization of high-frequency detail textures, but also ensure the analysis in low-frequency regions. In view of the varying importance of the feature vector, EWC distance assigns each element of a vector with different weight, which makes the low frequency components more discriminative [34]. Euclidean distance is one of the most widely used similarity measurement approaches for face recognition. In this paper, Euclidean distance is directly used for NSCTLBP features and EWC distance is chosen for Gabor features, a weighted measure rule is applied in the proposed method to integrate the advantages of NSCTLBP and Gabor features. The procedure is briefly described in Fig. 2.

Euclidean distance is explored to compute NSCTLBP feature distance.

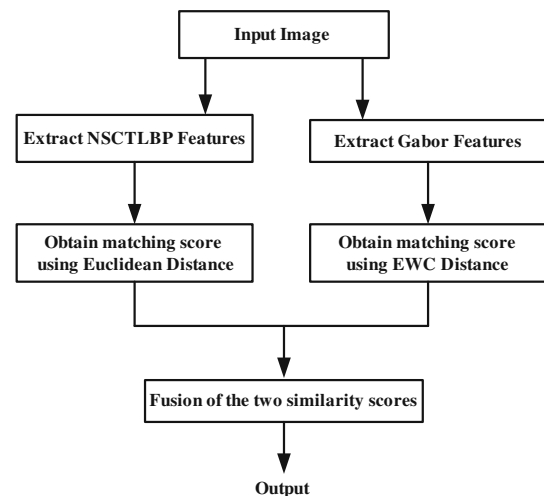


Fig. 2 Flow chart depicting the procedure involved in fusion of the NSCTLBP and Gabor features

$$D_{nsctlbp}(f_{nsctlbp}, f_{nsctlbp}^*) = \left(\sum_{p=1}^{u \times v} (f_{nsctlbp,p} - f_{nsctlbp,p}^*)^2 \right)^{\frac{1}{2}} \tag{8}$$

where $f_{nsctlbp}$ and $f_{nsctlbp}^*$ are NSCTLBP features of two face images to be matched respectively. $f_{nsctlbp,p}$ and $f_{nsctlbp,p}^*$ indicate the p -th components of $f_{nsctlbp}$, $f_{nsctlbp}^*$ separately.

According to EWC distance [35], Gabor feature distance is given by

$$D_{Gabor}(f_{Gabor}, f_{Gabor}^*) = \frac{\sum_{p=1}^{u \times v} ((\mathbf{W}_p \mathbf{W}_p^*) / \lambda_p^2)}{\left(\sum_{p=1}^{u \times v} (\mathbf{W}_p / \lambda_p)^2 \sum_{p=1}^{u \times v} (\mathbf{W}_p^* / \lambda_p)^2 \right)^{\frac{1}{2}}} \tag{9}$$

where f_{Gabor} and f_{Gabor}^* are Gabor features of two face images to be matched respectively. \mathbf{W}_p and \mathbf{W}_p^* indicate the p -th components of f_{Gabor} and f_{Gabor}^* separately. The eigenvalue denoted by λ_p can be calculated as follows.

Firstly, let $g = \{f_{Gabor,q}\}_{q=1}^Q$ denote a set of Q Gabor feature vectors extracted from Q training faces for one person in face database. The length of each feature vector is $u \times v$. The average facial feature of set g is computed by

$$\bar{f}_{Gabor} = \frac{1}{Q} \sum_{q=1}^Q f_{Gabor} \tag{10}$$

Secondly, calculate the covariance matrix of training faces $\xi = \mathbf{G}\mathbf{G}^T$, where \mathbf{G}^T is matrix transpose of \mathbf{G} .

$$\mathbf{G} = [f_{Gabor,1} - \bar{f}_{Gabor}, f_{Gabor,2} - \bar{f}_{Gabor}, \dots, f_{Gabor,Q} - \bar{f}_{Gabor}] \tag{11}$$

λ_p is the p -th eigenvalue of ξ .

Research shows that the measure principles of Euclidean distance and EWC distance are different. A lower $D_{nsctlbp}$ indicates high similarity of two features to be matched, especially $D_{nsctlbp} = 0$ denotes that they are identical. But the range of D_{Gabor} is in the interval $[-1, 1]$. $D_{Gabor} = 1$ indicates that the training face and the test face are identical.

Before fusion, normalization is required to map the matching scores obtained from multiple frameworks to a common range so that they can be easily combined. Thereafter, these individual matching distances are combined by using the sum rule to generate a single scalar score which is then used to make the final decision. Gabor feature matching score can be computed by

$$\tilde{D}_{Gabor} = (D_{Gabor} - 1) / 2 \tag{12}$$

where \tilde{D}_{Gabor} is in the interval $[-1, 0]$, $\tilde{D}_{Gabor} = 0$ indicates that the training face and the test face are identical.

Therefore on the basis of NSCTLBP feature matching score and Gabor feature matching score, a weighted

measure rule is proposed to measure the fusion of $\tilde{D}_{nsctlbp}$ and \tilde{D}_{Gabor} .

$$D = \omega \tilde{D}_{nsctlbp} + (1 - \omega)(-\tilde{D}_{Gabor}) \tag{13}$$

$$= \omega \tilde{D}_{nsctlbp} - (1 - \omega)\tilde{D}_{Gabor}$$

where $\tilde{D}_{nsctlbp}$ is NSCTLBP feature matching score, it is the normalized $D_{nsctlbp}$ to avoid the variations of large value of $D_{nsctlbp}$. ω is a weighted parameter ranging from 0 to 1. D is the matching score of the test face and training face. The smaller the value of D , the greater the similarity of the two faces, especially $D = 0$ denotes that they are identical.

4 Experimental Results and Analyses

4.1 Face Databases Explanation

In this paper typical Yale and ORL databases are adopted to verify the validity of the proposed method.

Table 1 gives the details about Yale and ORL databases. Figures 3 and 4 show the sample images for one person in each database. Both databases are divided into test subset and training subset. ORL-test subset consists of 200 images (5 images per person) and ORL-training subset contains the remaining 200 images. Yale-test subset contains 75 images (5 images per person), Yale-training subset contains the remaining 90 images(6 images per person).

4.2 Experiment Results and Analysis

In this paper, the ‘maxflat’ filter and ‘dmaxflat7’ filter are selected as NSP and NSDFB respectively in NSCT decomposition. We repeated 20 trails for random training and testing sets. Each face image is decomposed into three scales and each scale with eight orientations. Gabor feature extraction also complies in three different scales and samples in eight orientations for each scale, in addition, Gabor kernel parameters are set as follows: $\sigma = 2\pi$, $k_{max} = \pi/2$, $f = \sqrt{2}$.

- (1) Comparison on average face recognition rate of the methods based on different features

The proposed method in this paper fuses NSCTLBP feature and Gabor feature to measure the similarity between the test images and the training images. From Eq. (13), $\omega = 0$ indicates that only Gabor feature is considered for identification, while $\omega = 1$ means only NSCTLBP feature is considered. To illustrate the effectiveness of the proposed method, we make an average recognition rates comparison of the proposed method with the methods based on NSCTLBP feature, NSCT feature and Gabor feature separately on Yale and ORL databases.

Table 1 Description of ORL and Yale face databases

	Yale database	ORL database
Number of person	15	40
Number of images for one person	11	10
Size of database	15 × 11	40 × 10
Resolution	135 × 165	92 × 112
Interference	Expression and illumination variations, glasses occlusion	Expression and angle variations, glasses occlusion
Size of test subset	15 × 5	40 × 5
Size of training subset	15 × 6	40 × 5

**Fig. 3** Sample images for one person from Yale face database**Fig. 4** Sample images for one person from ORL face database

Table 2 shows that the proposed method with $\omega = 0.4$ is more effective than other three methods. The proposed method can extract facial information properly from both

Table 2 Average recognition rates of the methods based on different features

Face database	Methods based on different features			
	NSCT-feature (%)	NSCTLBP-feature (%)	Gabor-feature (%)	fusion feature ($\omega = 0.4$) (%)
Yale face	61.33	65.33	73.33	84.00
ORL face	59.50	75.00	77.50	86.50

low-frequency region and high-frequency region, which agrees with the above theoretical analysis.

(2) Anti-interference ability

For further research on the anti-interference performance of the proposed method, the test face subsets of Yale and ORL databases are classified in accordance with the interferences such as angle, expression, illumination variations and glasses occlusion. The average face recognition rates of different methods under different interferences are listed in Tables 3 and 4.

(a) Comparison on performance between NSCTLBP feature and NSCT feature

NSCT and NSCTLBP features have equivalent performance on Yale database under illumination variations and glasses occlusion on ORL database, but in general, NSCTLBP feature is more robust than NSCT feature on both Yale and ORL database, the reason is that LBP could capture some small texture details of face images. As expression and angle variations will cause partial rotation or translation of local facial textures, NSCTLBP feature with shift-invariant and rotation-invariant property would effectively adapt to these variations.

(b) Comparison on anti-interference between the proposed method and others

Recognition rates in Tables 3 and 4 indicate that the proposed method performs better than the other methods in the case of illumination, expression and angle variations. In addition to the rotation and translation of the facial textures, expression and angle variations will affect the distribution of both low-frequency region and high-frequency region of face image. Hence, the fusion texture feature is robust to expression and angle variations. Illumination variations primarily affects low-frequency region of face image. The fusion texture feature is superior to NSCTLBP feature and Gabor feature in describing high-frequency details alone. In the case of glasses occlusion, the proposed method contributes slightly to the recognition performance. That's because the glasses occlusion in the face images can severely obstruct the feature extraction in the eye area.

(3) The role of NSCTLBP and Gabor features in facial feature representation

Table 3 Average recognition rates of the methods under different interferences on Yale face database

Interference	Methods based on different features			
	NSCT-feature (%)	NSCTLBP-feature (%)	Gabor-feature (%)	fusion feature ($\omega = 0.4$) (%)
Illumination	53.33	53.33	40.00	60.00
Expression	68.89	71.11	80.00	91.11
Glasses	46.67	60.00	86.67	86.67

Table 4 Average recognition rates of the methods under different interferences on ORL face database

Interference	Methods based on different features			
	NSCT-feature (%)	NSCTLBP-feature (%)	Gabor-feature (%)	Fusion feature ($\omega = 0.4$) (%)
Angle	51.25	70.00	76.25	85.00
Expression	64.10	78.21	76.92	89.74
Glasses	72.22	72.22	77.78	77.78

Considering the role of NSCTLBP and Gabor features in facial feature representation, we take Yale database for example and make the following experiments by [36].

The vertical coordinate represents the matching score between the training image and the test images.

In this experiment, three test images with illumination, expression variations and glasses occlusion separately are selected from the same person to extract NSTPLBP and Gabor features. Figure 5 shows one of the examples of the training image and the test images with different interferences. Note that, Fig. 6 presents the average results of ten different persons from Yale database.

From Fig. 6, the matching score of Gabor feature (i.e. $-\hat{D}_{Gabor}$) and NSCTLBP feature (i.e. $\hat{D}_{nscctlbp}$) between training image and test images with illumination variations are larger than that with expression variations and glasses occlusion, which demonstrates that Gabor feature and NSCTLBP feature are more sensitive to illumination variations. The matching score of NSCTLBP feature is lower than that of Gabor feature, that is to say NSCTLBP feature is much robust to illumination variations than that of Gabor feature. In other words, in case of illumination variations the contribution of NSCTLBP feature is larger than that of Gabor feature. The same analyses are applicable to NSCTLBP and Gabor features in facial feature representation in case of glasses occlusion and expression variations. We can conclude that in case of expression variations the contribution of Gabor feature is larger and the contribution of NSCTLBP feature is larger in case of glasses occlusion.

(4) Performance of the proposed method with fewer training images for each person

In order to verify the effectiveness and adaptability of the proposed method with fewer training images,

experiments are performed on ORL and Yale databases by taking different number of training images randomly. The



Fig. 5 Example of the training image and the test images with different interferences. **a** Training image. **b** Test image with glasses occlusion. **c** Test image with expression variations. **d** Test image with illumination variations

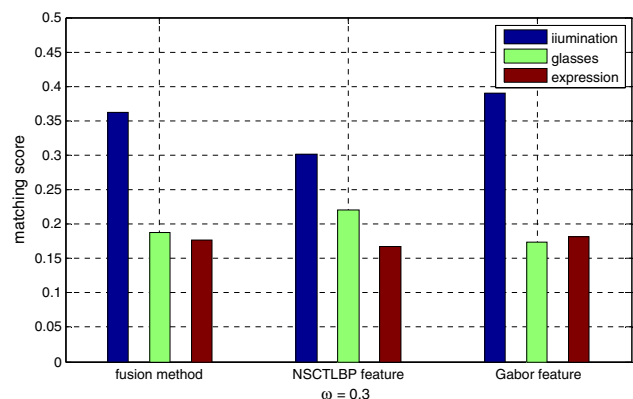


Fig. 6 Demonstration of different roles of NSCTLBP and Gabor features in face representation

average recognition rates over 20 different runs of training and test sets are presented in Tables 5, 6 and 7.

As the constructed NSCTLBP feature is invariant to rotation, and Gabor transform can extract almost all the desired features in high frequency and low frequency regions, the proposed method is especially robust to angle

Table 5 Average recognition rates of the proposed method with fewer training images

Face database	The number of training images for one person			
	2	3	4	5
Yale face	61.21	70.57	77.31	80.92
ORL face	74.28	80.66	83.78	86.50

and expression variations, thus the average recognition rate of the proposed method on ORL database is higher than that on Yale database. From the data obtained, it is observed that the recognition rates decline with fewer training images, because less desired features can be extracted with the decrease of training images. Additionally, illumination variations can affect distribution of the gray value of original face image. The average recognition rate of the proposed method with different training images in Table 6 also confirms the conclusion. With the variations of the training images of each person, exhaustive experiments testify the proposed method to be reasonable and robustness.

(5) Impact on recognition performance by ω

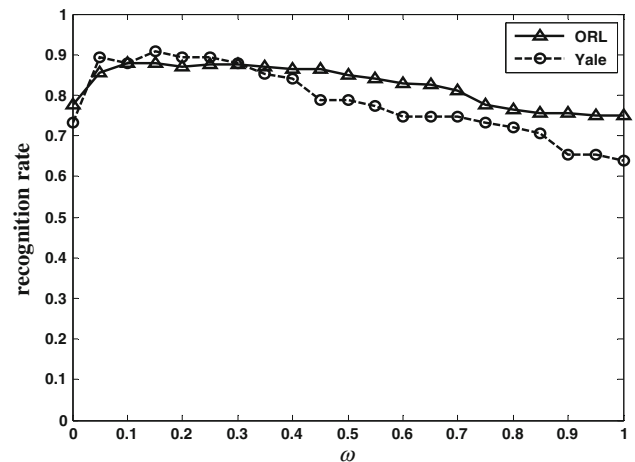
In order to examine the effect of the weighted parameter ω on the proposed method, experiments on ORL and Yale databases are implemented by ranging ω from 0 to 1. Figure 7 illustrates the results on Yale and ORL databases, where ω is taken as abscissa and recognition rate as vertical coordinate. Distribution of face recognition rate also reveals the recognition performance of the proposed method. Additionally, it can be seen from Fig. 7 that the method has a good recognition rate when ω ranges from 0.1 to 0.3, and the recognition rate can reach 88.89 % on ORL database and 93.33 % on Yale database when $\omega = 0.2$. When ω is out of the range, the recognition rate will decline. The reason is that Gabor feature plays a growing role in the proposed method as ω decreases, which leads to poor description of facial information in high-frequency region. While NSCTLBP feature plays a growing role in the proposed method as ω increases, which leads to less facial information in low-frequency region. We can conclude that the proposed method has the best

Table 6 Average recognition rates of the proposed method under different interferences on Yale

Interference	The number of training images for one person			
	2	3	4	5
Illumination	45.63	51.47	56.23	59.16
Expression	68.84	76.31	82.49	86.1
Glasses	59.03	61.46	76.51	86.16

Table 7 Average recognition rates of the proposed method under different interferences on ORL

Interference	The number of training images for each person			
	2	3	4	5
Angle	77.52	78.08	82	85.00
Expression	80.67	82.99	86.3	89.74
Glasses	64.34	66.75	72.53	77.78

**Fig. 7** Face recognition rate of the proposed method with different ω

performance by offering a good tradeoff between the constructed NSCTLBP feature and Gabor feature when ω ranges from 0.1 to 0.3.

5 Conclusions

A face recognition method fusing texture features is proposed and experiments on Yale and ORL databases are carried out to verify its recognition performance. Experimental results show that NSCTLBP feature constructed in this paper is invariant to rotation, and it is more robust than NSCT feature. Meanwhile the proposed method has better recognition performances than the methods based on NSCTLBP feature or Gabor feature. The proposed method not only retains the multi-scale and multi-orientation analysis capability of NSCT and Gabor transform but also takes advantage of NSCTLBP feature to compensate Gabor feature for the lack of facial information in high-frequency regions. Hence, it is robust to illumination, expression, angle variations and glasses occlusion.

Acknowledgments The authors gratefully acknowledge the editors and anonymous reviewers for their valuable comments. This paper is supported by Chongqing municipal education science twelfth five-year plan (2014-GX-055) and the science and technology research

project of Chongqing City Board of Education (KJ1402909 and KJ1503003).

References

- Eleyan A, Demirel H (2009) Co-occurrence based statistical approach for face recognition. In: 24th international symposium on computer and information sciences, pp 611–615
- Yu J, Li C (2013) Face recognition based on Euclidean distance and texture features. In: Proceedings of IEEE, international conference on computational and information sciences, pp 211–213
- Ojala T, Pietikäinen M, Mäenpää T (2002) Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans PAMI* 24:971–987
- Ahonen T, Hadid A, Pietikainen M (2006) Face description with local binary patterns: application to face recognition. *IEEE Trans PAMI* 28:2037–2041
- Ahonen T, Hadid A, Pietikainen M (2004) Face recognition with local binary patterns. *Comput Vis ECCV* 3021:469–481
- Zhang B, Gao Y, Zhao S, Liu J (2010) Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor. *IEEE Trans Image Process* 19:533–544
- Mehta R, Yuan J, Egiazarian K (2014) Face recognition using scale-adaptive directional and textural features. *Pattern Recogn* 47(5):1846–1858
- Daubechies I (1990) The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans Inf Theory* 36:961–1005
- Daugman JG (1985) Uncertainty relation for resolution in space, spatial-frequency, and orientation optimized by two-dimensional visual cortical filters. *J Opt Soc Am A* 2:1160–1169
- Daugman JG (1988) Complete discrete 2-d Gabor transforms by neural networks for image analysis and compression. *IEEE Trans Acoust Speech Signal Process* 36:1169–1179
- Xu Y, Li ZM, Pan JS et al (2013) Face recognition based on fusion of multi-resolution Gabor features. *Neural Comput Appl* 23(5):1251–1256
- Yang M, Zhang L, Shiu SCK et al (2013) Gabor feature based robust representation and classification for face recognition with Gabor occlusion dictionary. *Pattern Recogn* 46(7):1865–1878
- Zhang W, Shan S, Gao W, Chen X, Zhang H (2005) Local Gabor binary pattern histogram sequence (LGBPHS): a novel non-statistical model for face representation and recognition. In: Proceedings of international conference on computer vision, pp 786–791
- Zhang B, Shan S, Chen X, Gao W (2007) Histogram of Gabor phase patterns (HGPP): a novel object representation approach for face recognition. *IEEE Trans Image Process* 16:57–68
- Qin HF, Qin L, Xue L, Li YT (2012) A kernel Gabor-based weighted regions covariance matrix for face recognition. *Sensors* 12:7410–7422
- Xia W, Yin SY, Peng OY (2013) A high precision feature based on LBP and Gabor theory for face recognition. *Sensors* 13:4499–4513
- Bashyal S, Venayagamoorthy GK (2008) Recognition of facial expressions using Gabor wavelets and learning vector quantization. *Eng Appl Artif Intel* 21:1056–1064
- Tan XY, Triggs B (2010) Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans Image Process* 19:635–1650
- Yu L, He ZS, Cao Q (2010) Gabor texture representation method for face recognition using the Gamma and generalized Gaussian models. *Image Vis Comput* 28:177–187
- Srisuk S (2013) Robust face recognition based on texture analysis. *Int J Adv Rob Syst* 10:1–7
- Xie SF, Shan SG, Chen XL, Meng X, Gao W (2009) Learned local Gabor patterns for face representation and recognition. *Signal Process* 89:2333–2344
- Xie SF, Shan SG, Chen XL, Chen J (2010) Fusing local patterns of Gabor magnitude and phase for face recognition. *IEEE Trans Image Process* 19:1349–1361
- Zhang BC, Wang ZL, Zhong B (2008) Kernel learning of histogram of local Gabor phase patterns for face recognition. *EURASIP J Adv Signal Process*. doi:10.1155/2008/469109
- Han H, Zhu J, Lei Z et al (2013) Discriminant analysis with Gabor phase feature for robust face recognition. *J Electron Imaging* 22(4):043035
- Cunha AL, Zhou JP, Do MN (2006) The nonsubsampling contourlet transform: theory, design and applications. *IEEE Trans Image Process* 15:3089–3101
- Xu XB, Zhang DY, Zhang XM (2009) An efficient method for human face recognition using nonsubsampling contourlet transform and support vector machine. *Opt Appl* 39:601–615
- Wang Y (2013) Face recognition by nonsubsampling contourlet transform and support vector machine. In: Proceedings of the international conference on information engineering and applications, pp 595–600
- Xie XH, Lai JH, Zheng WS (2010) Extraction of illumination invariant facial features from a single image using nonsubsampling contourlet transform. *Pattern Recogn* 43:4177–4189
- Cheng Y, Hou Y, Zhao C, Li Z, Hu Y, Wang C (2010) Robust face recognition based on illumination invariant in nonsubsampling contourlet transform domain. *Neuro Comput* 73:2217–2224
- Cheng Y, Wang H, Liu HJ, Rong HJ (2009) Illumination normalization for robust face recognition based on nonsubsampling contourlet transform. *IEEE CCPR*, pp 1–5
- Fan CN (2012) Nonsubsampling contourlet transform based illumination invariant extracting method. *Adv Inform Sci Serv Sci* 4:47–55
- Do MN, Vetterli M (2005) The contourlet transform: an efficient directional multiresolution image representation. *IEEE Trans Image Process* 14:2091–2106
- Lee TS (1996) Image representation using 2D Gabor wavelets. *IEEE Trans PAMI* 18:959–971
- Štruc V, Pavešić N (2009) Gabor-based kernel partial-least-squares discrimination features for face recognition. *Informatica* 20(1):115–138
- Shi J, Samal A, Marx D (2006) How effective are landmarks and their geometry for face recognition. *Comput Vis Image Underst* 102:117–133
- Su Y, Shiguang S, Chen X, Gao W (2009) Hierarchical ensemble of global and local classifiers for face Recognition. *IEEE Trans Image Process* 18(8):1885–1896