

Image preprocessing method based on local approximation gradient with application to face recognition

Zhaokui Li¹ · Yan Wang¹ · Chunlong Fan¹ · Jinrong He²

Received: 30 March 2014 / Accepted: 5 March 2015 / Published online: 28 March 2015
© Springer-Verlag London 2015

Abstract In order to obtain more robust face recognition results, the paper proposes an image preprocessing method based on local approximation gradient (LAG). The traditional gradient is only calculated along 0° and 90° ; however, there exist many other directional gradients in an image block. To consider more directional gradients, we introduce a novel LAG operator. The LAG operator is actually calculated by integrating more directional gradients. Because of considering more directional gradients, LAG captures more edge information for each pixel of an image and finally generates an LAG image, which achieves a more robust image dissimilarity between images. An LAG image is normalized into an augmented feature vector using the “z-score” method. The dimensionality of the augmented feature vector is reduced by linear discriminant analysis to yield a low-dimensional feature vector. Experimental results show that the proposed method achieves more robust results in comparison with state-of-the-art methods in AR, Extended Yale B and CMU PIE face database.

Keywords Local approximate gradient · Image preprocessing · Linear discriminant analysis · Robust dissimilarity · Face recognition

1 Introduction

Face recognition (FR) is the problem of verifying or identifying a face from its image. It has received a great deal of attention from the scientific and industrial communities over the last three decades due to its wide range of applications in entertainment, smart cards, information security, law enforcement, access control and video surveillance [1, 2]. Various methods have been proposed for facial feature extraction and classification, among which the representatives include subspace learning methods (e.g., Eigenface [3], Fisherface [4], Laplacianfaces [5]), subspace learning from image gradient orientations (IGO) [6]), kernel-based subspace learning methods [7–10], local binary pattern (LBP) methods [11, 12], Gabor feature-based classification methods [13–15], preprocessing-based approaches [16–21, 30, 31, 33, 34], the recently developed sparse representation-based methods [22–25], and the like. While most of these methods in controlled conditions have already achieved impressive performance over large-scale galleries, there still exist many challenges for FR in uncontrolled environments, such as partial occlusions, varying lighting conditions, expressions, poses, and so on.

This paper focuses mainly on the image preprocessing technology, and the issue of robustness to facial expression varying, lighting variations and partial occlusions. Preprocessing-based approaches seek to reduce the image to a more “canonical” form, and many well-known algorithms have been developed to tackle this problem. Histogram equalization (HE) is one simple example, but purpose-

✉ Zhaokui Li
lmy52wy@163.com

Yan Wang
wy4615@sina.com

Chunlong Fan
fanchunlong@sau.edu.cn

Jinrong He
hejinrong@whu.edu.cn

¹ School of Computer, Shenyang Aerospace University, Daoyi South 37th Street, Daoyi economic development zone, Shenyang 110136, Liaoning, China

² College of Information Engineering, Northwest A&F University, Yangling 712100, Shanxi, China

designed methods often exploit the fact that (on the scale of a face) naturally occurring incoming illumination distributions typically have predominantly low spatial frequencies and soft edges so that high-frequency information in the image is predominantly signal (i.e., intrinsic facial appearance). For example, the Multiscale Retinex (MSR) method of Jobson et al. [16] cancels much of the low-frequency information by dividing the image by a smoothed version of itself. Zhang et al. developed a novel technique called Gradientfaces (GF) by extracting the illumination insensitive measure from the gradient domain [17]. Wang et al. [18] use a similar idea (with a different local filter) in the self-quotient image (SQI). Gross and Brajovic (GB) [19] developed an anisotropic smoothing method that relies on the iterative estimation of a blurred version of the original image. More recently, Xiaoyang Tan and Bill Triggs (TT) [20] present a simple and efficient preprocessing chain that eliminates most of the effects of changing illumination while still preserving the essential appearance details that are needed for recognition, and B. Wang et al. [21] present a novel Illumination normalization technique called Weberfaces (WF) by using Weber's law. In addition, nonnegative matrix factorization (NMF) has become a popular data-representation method and has been widely used in image processing and pattern-recognition problems [30, 31].

In this paper, we propose an image preprocessing method based on local approximation gradient (LAG). The traditional gradient is only calculated along 0° and 90° ; however, there exist many other directional gradients in an image block. In order to consider more directional gradients, we introduce a novel LAG operator. The LAG operator is actually calculated by integrating more directional gradients. Generally, the main purpose of gradient calculation is to obtain different edges. Because of considering more directional gradients, LAG captures more edge information for each pixel of an image and finally generates an LAG image, which achieves a more robust image dissimilarity between images. More edge information depicts higher variances of the image. Intuitively, one often pays more attention to the regions of higher variances in a given image as compared with the flat regions. Therefore, the computation of orientation LAG should play an important role in a classification task. Then, a LAG image is normalized into an augmented feature vector using the "z-score" method. The dimensionality of the augmented feature vector is reduced by linear discriminant analysis to yield a low-dimensional feature vector. Experimental results show that LAG achieves more robust results in comparison with state-of-the-art methods in AR, Extended Yale B and CMU PIE face database. The rest of the paper is organized as follows: Section 2 presents our preprocessing method, Sect. 3 reports experimental results, and Sect. 4 concludes the paper.

2 Image preprocessing method based on local approximation gradient

2.1 Local approximation gradient

Generally, gradient operator is a first-order derivative operator. For an input image $\Gamma(x, y)$, its gradient is a vector:

$$\nabla\Gamma(x, y) = \begin{pmatrix} G_x \\ G_y \end{pmatrix} = \begin{pmatrix} \partial\Gamma & \partial x \\ \partial\Gamma & \partial y \end{pmatrix}. \quad (1)$$

The amplitude of the vector is expressed as:

$$\text{mag}(\Gamma) = (G_x^2 + G_y^2)^{1/2} \quad (2)$$

The orientation angle of the vector is expressed as:

$$\beta(x, y) = \arctan\left(\frac{G_y}{G_x}\right) \quad (3)$$

For a discrete image $\Gamma(x, y)$, G_x and G_y are usually calculated by different templates. An image template neighborhood $T \in \mathfrak{R}^{3 \times 3}$ is shown in Fig. 1. We denote by α the intensity value in the central pixel location and by $b_m (m = 1, 2, \dots, 8)$ the m th neighbor of α . The orientation angle of α can be calculated according to $\beta_\alpha = \arctan\left(\frac{b_3 - b_7}{b_1 - b_5}\right)$.

We can see that the traditional gradient is only calculated along 0° and 90° ; however, there exist many other directional gradients in an image block. To consider more directional gradients, we design a novel gradient operator called LAG. The local approximate gradient should also be calculated in a local image block. Figure 2 shows eight different directions in 3×3 image template neighborhood.

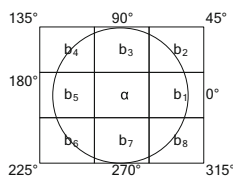
For an input image $\Gamma(x, y)$, we can calculate its LAG vector according to formula (4).

$$\nabla\Gamma^{\text{LAG}}(x, y) = \begin{pmatrix} G_{0^\circ} \\ G_{45^\circ} \\ G_{90^\circ} \\ G_{135^\circ} \\ G_{225^\circ} \\ G_{270^\circ} \\ G_{315^\circ} \end{pmatrix} = \begin{pmatrix} (b_1 - \alpha) \\ (b_2 - \alpha) \\ (b_3 - \alpha) \\ (b_4 - \alpha) \\ (b_5 - \alpha) \\ (b_6 - \alpha) \\ (b_7 - \alpha) \\ (b_8 - \alpha) \end{pmatrix} \quad (4)$$

Fig. 1 Image template neighborhood

b_4	b_3	b_2
b_5	α	b_1
b_6	b_7	b_8

Fig. 2 Eight different directions in image template neighborhood



The amplitude of the vector is expressed as:

$$\begin{aligned} \text{mag}(\Gamma^{\text{LAG}}) &= (G_{0^\circ}^2 + G_{45^\circ}^2 + \dots + G_{315^\circ}^2)^{1/2} \\ &= \left(\sum_{m=1}^8 (b_m - \alpha)^2 \right)^{1/2} \end{aligned} \tag{5}$$

The orientation angle of the vector is expressed as:

$$\beta^{\text{LAG}} = \arctan \left(\frac{1}{8} \left(\sum_{m=1}^8 b_m - \alpha \right) \right) \tag{6}$$

Suppose there are N pixels in an image. We treat a pixel α in an image as a center and determine its l neighbor pixels on a square of radius L using the city-block distance (note that for the pixel on the margin of an image, we use the mirror transform to enlarge the image first and then determine its neighbors). These neighbor pixels form a squarely neighbor set. We denote by $\Omega_\alpha^L = \{b_1, b_2, \dots, b_l\}$, $l = \#(\Omega_\alpha^L)$ the L th neighbor set of pixel α . L is radius, and l is the number of neighbor pixel.

Figure 3 illustrates different neighbor sets for different radius L . Figure 3a shows the first neighbor set Ω_α^1 (including 8 neighbor pixels). Figure 3b shows the second neighbor set Ω_α^2 (including 24 neighbor pixels). Figure 3c shows the second neighbor set Ω_α^3 (including 48 neighbor pixels).

For the L th neighbor set of pixel α , we can calculate the corresponding orientation angle according to formula (7)

$$\begin{aligned} \beta^{\text{LAG}} &= \arctan \left(\frac{1}{l} \left(\sum_{m=1}^l b_m - \alpha \right) \right), \Omega_\alpha \\ &= \{b_1, b_2, \dots, b_l\}, l = \#(\Omega_\alpha) \end{aligned} \tag{7}$$

For each pixel of an image, we consider all neighbor pixels in a local block, and calculate more directional gradients. The LAG is actually calculated by integrating more directional gradients. Generally, the main purpose of gradient calculation is to obtain different edge information. Because of considering more directional gradients, LAG should obtain more edge information. More edge information depicts higher variances of the image. Intuitively, one often pays more attention to the regions of higher variances in a given image compared with the flat regions. So, the computation of orientation LAG should play an important role in a classification task.

In the following, we will compare the difference dissimilarity between the different preprocessed images by different distance metric methods. Let us consider a motivating example in which different dissimilarity measures are applied to the images shown in Fig. 4. Figure 4 shows the different preprocessed images (including Original images, HE images, GF images, GB images, MSR images, TT images, WF images, and LAG images), and the resolution of all images resolution is adjusted to 60×60 . For GB [19] we set $\lambda = 1$. For MSR [16] we set $\text{hsiz} = [7, 15, 21]$. For GF [17] we set $\sigma = 0.75$. For TT [20] we set $\gamma = 0.2$, $\sigma_0 = 1$, $\sigma_1 = 2$, $\alpha = 0.1$, and $\tau = 10$. For WF [21] we set $\sigma = 0.75$, $\text{nn} = 9$, and $\alpha = 2$. For the proposed LAG, we set $L = 2$. Table 1 shows the dissimilarity comparison between neural expression image and the other images with different measures. The other images consist of smile expression image, sunglasses image, scarves image from same subject and neural expression images from different subject. These measure methods consist of Euclidean, Cosine, Correlation and Manhattan [29]. As can be seen in Table 1, seven methods (including None, HE, GF, GB, MSR, TT and WF) associate a smaller distance between the original neural image and a neural image from a different subject. The distance between the original and the same image with occlusion is larger. In contrast, the usage of LAG results in a large distance between the original neural expression image and the image of a different

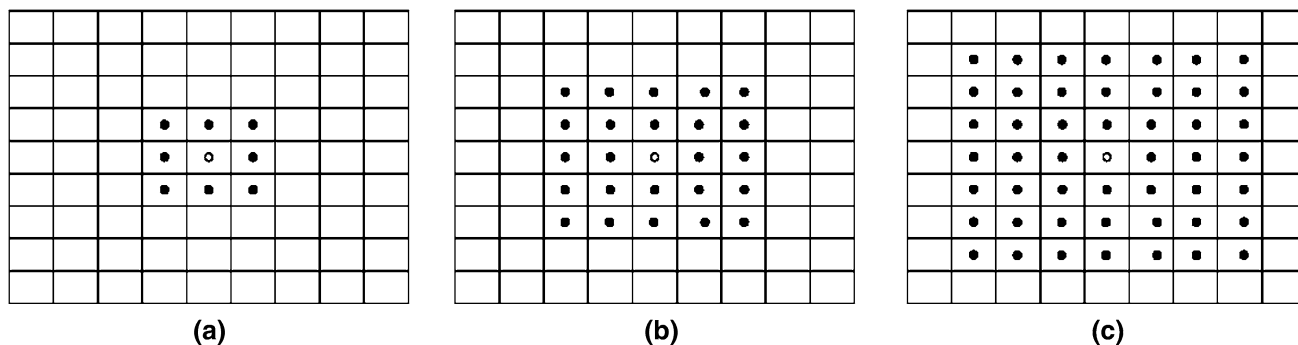


Fig. 3 Squarely neighbor sets for radius L . **a** $l = 8$ ($L = 1$), **b** $l = 24$ ($L = 2$), and **c** $l = 48$ ($L = 3$)

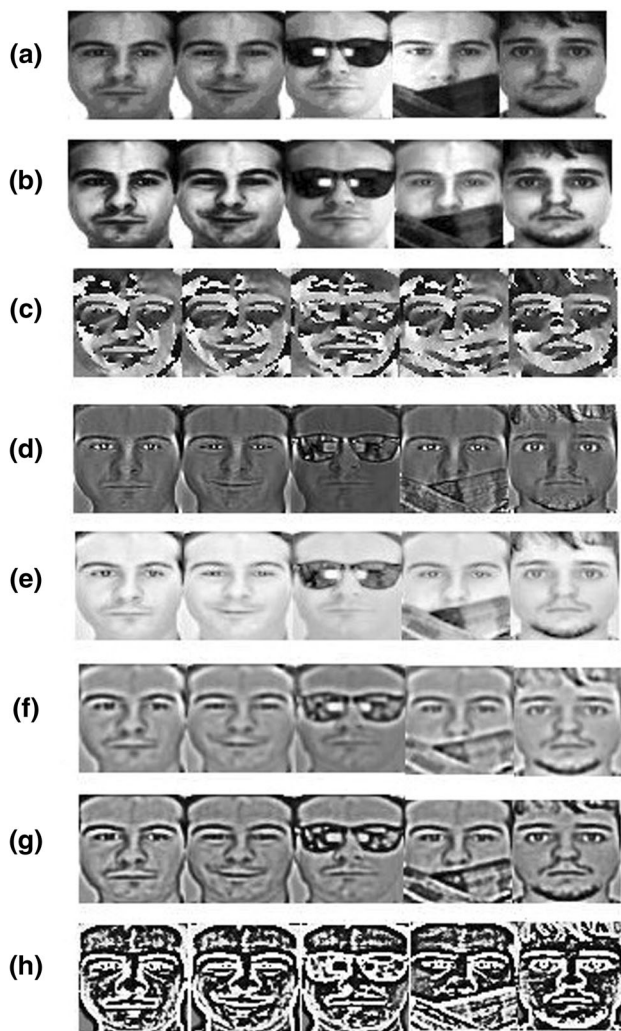


Fig. 4 Images used for the dissimilarity measurement. **a** Original images, **b** HE images, **c** GF images, **d** GB images, **e** MSR images, **f** TT images, **g** WF images, **h** LAG images. Each row from left to right in turn is the neutral expression, smiling expression, occluded (*sunglasses*), occluded (*scarves*) and another person's neural expression

person. Therefore, we can believe that LAG obtains more robust dissimilarity measure.

2.2 Image normalization

Assume that we are given a set of n input images $\{\Gamma_i\}$, $i = 1, 2, \dots, n$, we can obtain its local approximate gradient image $\{\Psi_i\}$, $i = 1, 2, \dots, n$ by calculating local approximate gradient. Let $\{y_i\}$ be the p -dimensional column vector obtained by writing $\{\Psi_i\}$ in lexicographic ordering. All column vectors form a matrix $B = [y_1 \dots y_n] \in \mathbb{R}^{p \times n}$. Each gradient image is normalized using the “z-score” method [32]. We denote by δ_i the normalized vector of y_i . The q th dimension of δ_i can be calculated according to formula (8).

$$\delta_i^{(q)} = \left(y_i^{(q)} - \mu^{(q)} \right) / \sigma^{(q)}, \quad i = 1, \dots, n, \quad q = 1, \dots, p \quad (8)$$

where $y_i^{(q)}$, $\delta_i^{(q)}$ are, respectively, the q th dimension of column vector y_i and δ_i , $\mu^{(q)}$ and $\sigma^{(q)}$ are the corresponding mean and variance respectively, and they can be calculated according to $\mu^{(q)} = \frac{1}{n} \sum_{i=1}^n y_i^{(q)}$ and $\sigma^{(q)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i^{(q)} - \mu^{(q)})^2}$. We can compute all p dimensions of column vector δ_i according to formula (8). For an input image Γ_i , we can obtain its corresponding normalized vector δ_i .

2.3 Dimensionality reduction

The dimensionality of the normalized vector is reduced by LDA to yield a low-dimensional feature vector. LDA searches for feature vectors in the underlying space that best discriminate among classes. More formally, given a number of independent features relative to which the data is described, LDA creates a linear combination of these which yields the largest mean differences between the desired classes. Mathematically speaking, for all the samples of all classes, we define two measures: (1) one is called within-class scatter matrix, as given by formula (9)

$$S_w = \sum_{c=1}^C \sum_{h=1}^{l^c} (\delta_{h,c} - m^c)(\delta_{h,c} - m^c)^T \quad (9)$$

where $\delta_{h,c}$ denotes the h th training sample in class c , l^c is the number of training samples in class c , m^c is the mean of the training samples in class c , and C is the number of class. (2) The other is called between-class scatter matrix, as given by formula (10)

$$S_b = \sum_{c=1}^C l^c (m^c - m)(m^c - m)^T \quad (10)$$

where m represents the mean of all classes.

The goal is to maximize the between-class measure while minimizing the within-class measure. The discriminative projection basis W and eigen values Λ are obtained by solving the thin linear discriminant analysis (LDA) eigendecomposition:

$$(S_w + \varepsilon I)^{-1} S_b W = W \Lambda \quad (11)$$

where ε is a small regularization constant (10^{-3} below) and I is the identity matrix. Thus, the original L -dimensional space is projected onto a final d -dimensional space using LDA.

Table 1 Dissimilarity comparison between neural expression image and the other images with different measures

Measure	Image	None	HE	GF	GB	MSR	TT	WF	LAG
Euclidean	Smile	0.1404	0.2264	0.5147	0.2621	0.1090	0.1593	0.2595	1.0181
	Sunglasses	0.5153	0.5363	0.6150	0.4189	0.2513	0.3188	0.425	1.0859
	Scarves	0.3912	0.6090	0.5921	0.2707	0.2084	0.2015	0.343	1.1507
	Different	0.3283	0.4653	0.6116	0.3654	0.2193	0.2691	0.3166	1.2571
Cosine	Smile	0.0943	0.0256	0.1325	0.0344	0.0059	0.0127	0.0337	0.5412
	Sunglasses	0.5419	0.1438	0.1891	0.0877	0.0316	0.0508	0.0903	0.6898
	Scarves	0.7851	0.1855	0.1753	0.0366	0.0217	0.0203	0.0588	0.6160
	Different	0.3927	0.1082	0.1832	0.0668	0.0240	0.0362	0.0501	0.8258
Correlation	Smile	0.0931	0.0996	0.5207	0.3868	0.2026	0.2499	0.3018	0.5306
	Sunglasses	0.5405	0.5586	0.6674	0.8451	0.6886	0.7359	0.6665	0.7207
	Scarves	0.7857	0.7226	0.7168	0.4381	0.4692	0.4063	0.4620	0.6544
	Different	0.3917	0.4213	0.7087	0.7207	0.6126	0.6856	0.5114	0.8710
Manhattan	Smile	5.1683	8.5085	17.2237	8.6717	3.5439	6.1040	9.7176	38.4434
	Sunglasses	17.6365	24.1771	24.9745	15.1932	9.5082	12.4320	16.8066	49.5127
	Scarves	22.6849	28.1592	24.4500	10.7663	8.6399	8.7246	14.1940	46.3281
	Different	14.4169	20.4238	24.7419	14.3901	8.5778	11.8163	13.1678	57.9166

Bold values indicate the maximum distance in certain distance measure

2.4 Robustness to occlusion

In order to evaluate the robustness to occlusion of our algorithm, we considered a single-sample-per-class experiment using neutral face images taken from the AR database. Our training test consisted of 100 face images of 100 different subjects, taken from session 1. Our testing set consisted of one image per subject taken from session 2. We evaluated the performance of our algorithms for the case of synthetic occlusions. Our algorithms consisted of LAG, and LAG with LDA (LAG-LDA). All test images were artificially occluded by a dinosaur patch of increasing size at random location. Figure 5 shows the corresponding corrupted images. Figure 6 shows the recognition rates of different algorithms without LDA. Figure 7 shows the recognition rates of different algorithms with LDA. As shown in Figs. 6 and 7, our algorithms (including LAG and LAG-LDA) consistently outperform the compared methods. As we may observe, LAG and LAG-LDA obtain more robust performance with a recognition rate over 75 % even when the percentage of occlusion is about 4 %. In addition, LAG achieves also better accuracy on FR than the compared methods without LDA and with LDA.

Figure 8 shows the recognition rates of LAG and LAG-LDA. As shown in Fig. 8, the two results are approximate.



Fig. 5 Artificial occlusion images using a dinosaur patch of increasing size at random location

Figure 9 shows the running time of all methods (including with LDA and without LDA). It’s worth noting that the methods with LDA achieves faster running time than that without LDA.

3 Experiments

In this section, we evaluate the robustness of our method for the application on FR. Three publicly available face databases, namely, AR [26], Extended Yale B [27], and CMU PIE [28] are used for experimental evaluation. For AR, Extended Yale B and CMU PIE, we resize them to

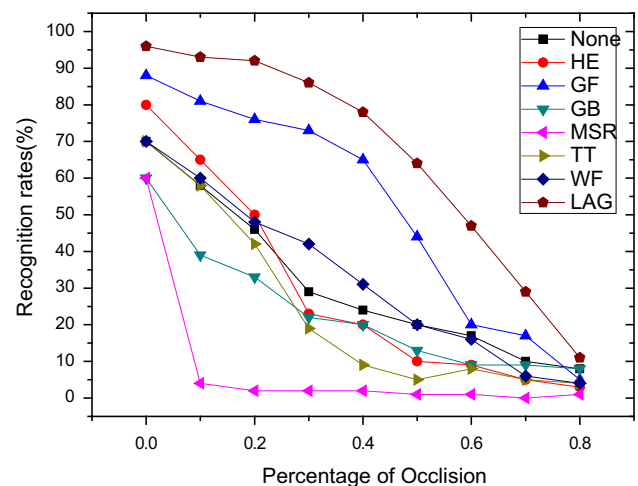


Fig. 6 Recognition rate as a function of the percentage of occlusion (without LDA)

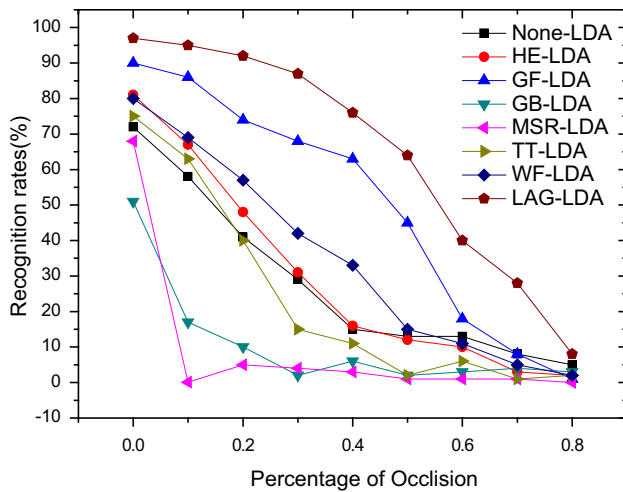


Fig. 7 Recognition rate as a function of the percentage of occlusion (with LDA)

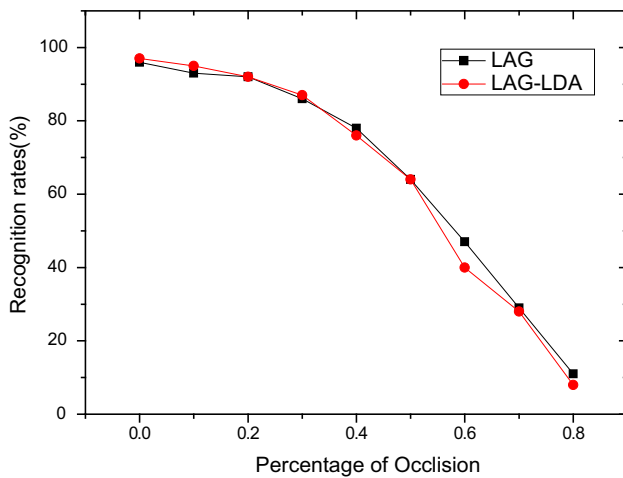


Fig. 8 Recognition rate as a function of the percentage of occlusion (LGA and LAG-LDA)

60 × 60. To verify the effectiveness of the proposed method, in this series of experiments we evaluate the performance of the proposed method and compare it to that of several methods using the LDA and 1-NN classifier. In order to evaluate the performance of our algorithm for the case of synthetic occlusions, a series of similar experiments were also performed. In these experiments, all test images were artificially occluded by a dinosaur patch at random position. We used dinosaur patch of resolution 30 × 30.

3.1 Parameter selection

For GB [19] we set $\lambda = 1$ for all experiments. For MSR [16] we set $hsiz = [7, 15, 21]$. For GF [17] we set $\sigma = 0.75$. For TT [20] we set $\gamma = 0.2, \sigma_0 = 1, \sigma_1 = 2, \alpha = 0.1,$ and $\tau = 10$. For WF [21] we set $\sigma = 0.75,$

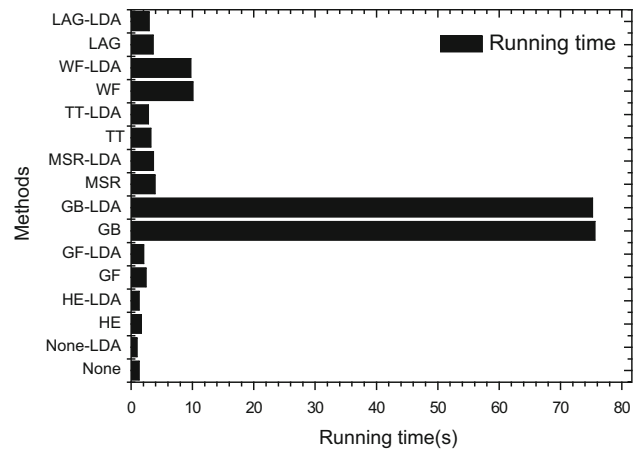


Fig. 9 Running time as a function of the percentage of occlusion

$nn = 9,$ and $\alpha = 2$. For the proposed LAG, we set the neighbor radius $L = 2$ for AR and Extended Yale B. In CMU PIE, the neighbor radius is set to 4. In addition, in 1-NN classifier, the dissimilarity measure based on cosine distance is adopted. In LDA, PCA ratio is set to 1, so all the non-zero eigenvalues will be kept.

3.2 AR database

The AR database consists of more than 4000 frontal view face images of 126 subjects. Each subject has up to 26 images taken in two sessions. The first session contains 13 images, numbered from 1 to 13, including different facial expressions (1–4), illumination changes (5–7), and different occlusions under different illumination changes (8–13). The second session duplicates the first session 2 weeks later. We randomly select a subset with 100 subjects.

Figure 10 shows a sample of images used in our experiments. Figure 10a shows the non-occluded images of session 1. Figure 10b shows the non-occluded images of session 2. Figure 10c shows the corresponding artificial occlusion images of (b). Figure 10d shows the face images occluded by the scarf or sunglasses.

We investigate the robustness of our scheme for the case of facial expression varying, illumination variations, occlusions(sunglasses)-illumination changes, and occlusions(scarves)-illumination changes. Table 2 provides the detailed information of each experiment (referred to as Exp.). The best recognition rates are shown in Table 3, Tables 4 and 5. The dimension that results in the best performance for each method is given in parentheses.

From Exp. 1 to Exp. 4, we use images 1–4 of session 1 for training. As can be seen from Table 3, LAG-LDA achieves better recognition rate compared with other methods. LAG-LDA achieves 93.25 % recognition rate for different expressions (Exp. 1). For Exp. 2, LAG-LDA

Fig. 10 Face images of the same subject taken from the AR database. **a** Shows the non-occluded images of session 1. **b** Shows the non-occluded images of session 2. **c** Shows the corresponding artificial occlusion images of **b**. **d** Shows the face images occluded by the scarf or sunglasses



Table 2 Detailed information of all experiments of the AR database

Exp.	Training set (session 1)	Test set (session 2)
Exp. 1	1–4	1–4 expressions (neural, smile, anger and scream)
Exp. 2	1–4	5–7 illumination variations
Exp. 3	1–4	8–10 sunglasses-illumination
Exp. 4	1–4	11–13 scarves-illumination
Exp. 5	8–13	1–4 expressions (neural, smile, anger and scream)
Exp. 6	8–13	5–7 illumination variations
Exp. 7	8–13	8–10 sunglasses-illumination
Exp. 8	8–13	11–13 scarves-illumination
Exp. 9	1–4	1–4 expressions (neural, smile, anger and scream) + dinosaur occlusion
Exp. 10	1–4	5–7 illumination variations + dinosaur occlusion
Exp. 11	8–13	1–4 expressions (neural, smile, anger and scream) + dinosaur occlusion
Exp. 12	8–13	5–7 illumination variations + dinosaur occlusion

achieves 89.33 % recognition rate for the case of illumination changes. For Exp. 3 and Exp. 4, which are more difficult experiment on the AR database, LAG-LDA achieves 67 % recognition rate for the case of occlusions(sunglasses)-illumination changes, and achieves 67.33 % recognition rate for the case of occlusions (scarves)-illumination changes.

Furthermore, from Exp. 5 to Exp. 8, we use images 8–13 of session 1 for training, and they are also very difficult experiment on the AR database. As can be seen from Table 4, LAG-LDA achieves also better recognition rate compared with other methods. LAG-LDA achieves 75.25 % recognition rate for different expressions (Exp. 5). For Exp. 6, LAG-LDA achieves 90.67 % recognition rate for the case of illumination changes. Due to the use of the occlusion images as the training set, in Exp. 7, LAG-LDA has also achieved better results, and achieves 92 % recognition rate for the case of occlusions(sunglasses)-

illumination changes. For Exp. 8, LAG-LDA achieves 73.33 % recognition rate for the case of occlusions (scarves)-illumination changes.

In order to evaluate the performance of our algorithm for the case of synthetic occlusions, we carried out four additional experiments (including Exp. 9, Exp. 10, Exp. 11 and Exp. 12) in AR database. Exp. 9, Exp. 10, Exp. 11 and Exp. 12 are similar experiments to Exp. 1, Exp. 2, Exp. 5 and Exp. 6. The only difference is that all test images were artificially occluded by a dinosaur patch at random position. Table 5 shows the corresponding results. As can be seen from Table 5, the proposed LAG-LDA performs better than other state-of-the-art methods. The performances of these compared methods are not satisfying. It appears that those compared seven methods are sensitive to an artificial occlusion image. In contrast, LAG-LDA is more robust to artificial occlusion image and thus achieves better results.

Table 3 Best recognition rate comparison on AR face database with different methods (from Exp. 1 to Exp. 4)

Method	Exp. 1	Exp. 2	Exp. 3	Exp. 4
None-LDA	88 % (89)	51.33 % (97)	16 % (85)	34 % (92)
HE-LDA	89 % (97)	53.33 % (99)	19 % (95)	20 % (95)
GF-LDA	88.25 % (99)	82 % (99)	39 % (62)	45.33 (99)
GB-LDA	61.5 % (93)	52.33 % (99)	20.33 % (97)	36.33 % (97)
MSR-LDA	71.75 % (87)	68.3 % (92)	12 % (99)	27 % (87)
TT-LDA	76.25 % (92)	73.67 % (95)	25.67 % (96)	50 % (99)
WF-LDA	68.5 % (99)	62.67 % (99)	31.67 % (99)	35.67 % (86)
LAG-LDA	93.25 % (91)	89.33 % (93)	67 % (96)	67.33 % (99)

Bold values indicate the best recognition rate

Table 4 Best recognition rate comparison on AR face database with different methods (from Exp. 5 to Exp. 8 and with LDA)

Method	Exp. 5	Exp. 6	Exp. 7	Exp. 8
None-LDA	57.25 % (92)	80 % (97)	78.67 % (89)	65.67 % (96)
HE-LDA	58 % (97)	77.33 % (99)	77.67 % (98)	53.33 % (98)
GF-LDA	64 % (97)	87.33 % (93)	86 % (97)	60 % (92)
GB-LDA	37 % (99)	41.67 % (98)	50 % (91)	37.67 % (85)
MSR-LDA	46.5 % (99)	55.67 % (98)	67.33 % (98)	39 % (81)
TT-LDA	50.5 % (99)	68 % (99)	66 % (94)	50.33 % (99)
WF-LDA	40.25 % (90)	51.67 % (91)	59 % (99)	36.67 % (96)
LAG-LDA	75.25 % (97)	90.67 % (95)	92 % (84)	73.33 % (99)

Bold values indicate the best recognition rate

Table 5 Best recognition rate comparison on AR face database with different methods (from Exp. 9 to Exp. 12)

Method	Exp. 9	Exp. 10	Exp. 11	Exp. 12
None-LDA	26.5 % (98)	11.67 % (98)	17.5 % (88)	22.33 (96)
HE-LDA	30.5 % (97)	11 % (46)	17 % (92)	25 % (85)
GF-LDA	66.75 % (95)	48.67 % (92)	38.25 % (98)	54 % (97)
GB-LDA	14.75 % (98)	11 % (93)	8.75 % (74)	11 % (74)
MSR-LDA	10 % (85)	6 % (75)	6.25 % (64)	8 % (51)
TT-LDA	10.5 % (97)	7.3 % (98)	7.5 % (97)	8.3 % (94)
WF-LDA	25.2 % (97)	17 % (70)	13.5 % (98)	17 % (98)
LAG-LDA	80.5 % (97)	75.67 % (98)	61.75 % (97)	75 % (97)

Bold values indicate the best recognition rate

3.3 Extended Yale B database

The Extended Yale B database contains 2414 front-view face images of 38 individuals. For each individual, about 64 pictures were taken under various laboratory-controlled lighting conditions. Figure 11 shows partial face images used in our experiments. Figure 11a shows the non-occluded images of one subject. Figure 11b shows the corresponding artificial occlusion images of (a).

We carried out the following two experiments in Extended Yale B database. In Exp. 1, a random subset with $d(=15, 20, 25, 32)$ images per individual is taken with labels to form the training set, and the rest of the database is

considered to be the testing set. We summarize the best recognition rates of these methods in Table 6. The dimension that results in the best performance for each method is given in parentheses. In Exp. 2, we performed a similar experiment with the Exp. 1, and the only difference is that all test images were artificially occluded by a dinosaur patch at random position. We summarize the best recognition rates of these methods in Table 7. The dimension that results in the best performance for each method is given in parentheses.

As can be seen from Table 6, LAG-LDA obtains better recognition rate compared to the other algorithms. In all experiments, LAG-LDA has more than 99 % recognition



Fig. 11 Partial face images of the same subject taken from the Extended Yale B database. **a** Shows the non-occluded images of one subject, and the corresponding images were artificially occluded by a dinosaur patch at random position in **(b)**

Table 6 Best recognition rate comparison on Extended Yale B face database with different methods

Method	Exp. 1				
	15	20	25	32	Std.
None-LDA	90.89 % (37)	88.82 % (34)	92.01 % (37)	89.73 % (37)	0.0139
HE-LDA	46.26 % (37)	43.11 % (37)	39.55 % (36)	40.65 % (35)	0.0298
GF-LDA	99.51 % (36)	99.7 % (32)	99.86 % (35)	99.92 % (21)	0.0018
GB-LDA	99.57 % (37)	99.82 % (35)	99.8 % (30)	100 % (19)	0.0018
MSR-LDA	99.95 % (29)	100 % (28)	99.93 % (31)	100 % (14)	0.0004
TT-LDA	99.89 % (22)	100 % (24)	100 % (31)	100 % (18)	0.0006
WF-LDA	99.78 % (34)	99.4 % (36)	99.8 % (31)	100 % (17)	0.0025
LAG-LDA	99.95 % (30)	100 % (32)	100 % (37)	100 % (18)	0.0003

Bold values indicate the best recognition rate

Table 7 Best recognition rate comparison on Extended Yale B face database with different methods (dinosaur occlusion)

Method	Exp. 2				
	15	20	25	32	Std.
None-LDA	78.69 % (34)	81.62 % (37)	83.74 % (37)	86.39 % (37)	3.2615
HE-LDA	26.46 % (37)	25.39 % (37)	27.39 % (37)	30.26 % (36)	13.8527
GF-LDA	95.98 % (37)	96.92 % (37)	97.61 % (37)	98.58 % (37)	1.0982
GB-LDA	91.90 % (37)	93.22 % (37)	93.78 % (37)	95.47 % (37)	1.4791
MSR-LDA	86.75 % (37)	89.32 % (37)	90.81 % (37)	90.94 % (37)	1.9473
TT-LDA	75.09 % (37)	70.63 % (37)	78.09 % (37)	78.13 % (37)	3.5360
WF-LDA	94.62 % (37)	94.79 % (37)	96.24 % (36)	97.41 % (37)	1.3157
LAG-LDA	99.40 % (37)	99.46 % (26)	99.59 % (36)	99.66 % (31)	0.1187

Bold values indicate the best recognition rate

rate. For each method, we calculate the standard deviation for all the recognition rates. As can be seen from Table 6, LAG-LDA obtains smaller standard deviations compared to the other methods, so the performance of our algorithm is more stable. Therefore, for illumination changes, our algorithm is more robust.

As can be seen from Table 7, we can see that the performances of all methods degrade compared with Table 6. However, LAG-LDA still obtains a better recognition rate as compared to the other algorithms. In all experiments, LAG-LDA has also more than 99 % recognition rate. For each method, we also calculate the standard deviation for all the recognition rates. As can be seen from Table 7, LAG-LDA

still obtains a smaller standard deviation as compared to the other methods, so the performance of our algorithm is more stable. Therefore, we can believe that our algorithm is more robust to illumination changes with different occlusions.

3.4 CMU PIE database

The CMU PIE database consists of more than 41,000 face images of 68 subjects. The database contains faces under varying pose, illumination, and expression. We used the five near frontal poses (C05, C07, C09, C27, C29) and a total of 170 images for each subject. Figure 12 shows partial face images used in our experiments. Figure 12a

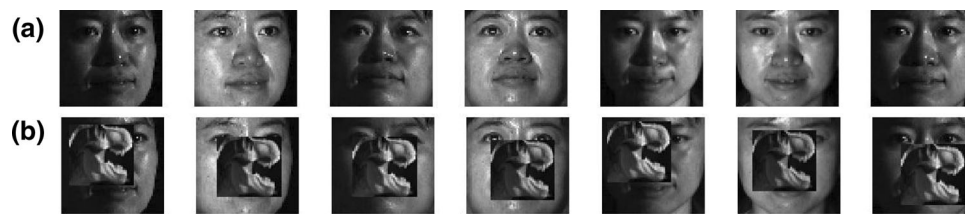


Fig. 12 Partial face images of the same subject taken from the CMU PIE database. **a** Shows the non-occluded images of one subject, and the corresponding images were artificially occluded by a dinosaur patch at random position in **(b)**

Table 8 Best recognition rate comparison on Extended Yale B face database with different methods

Method	Exp. 1					Std.
	10	15	20	25		
None-LDA	82.98 % (67)	86.22 % (67)	88.73 % (66)	88.88 % (59)	2.7654	
HE-LDA	8.9 % (55)	10.34 % (61)	10.78 % (66)	11.41 % (64)	1.0663	
GF-LDA	82.21 % (66)	86.69 % (67)	89.41 % (67)	89.82 % (67)	3.5022	
GB-LDA	81.80 % (66)	87.02 % (67)	88.28 % (67)	90.18 % (67)	3.5899	
MSR-LDA	86.58 % (66)	91.14 % (66)	91.23 % (64)	91.82 % (65)	2.4271	
TT-LDA	84.63 % (67)	88.61 % (67)	89.61 % (67)	90.12 % (66)	2.4887	
WF-LDA	86.07 % (67)	89.79 % (67)	90.88 % (67)	91.65 % (66)	2.4724	
LAG-LDA	92.35 % (66)	95.17 % (63)	96.13 % (57)	96.71 % (62)	1.9339	

Bold values indicate the best recognition rate

shows the non-occluded images of one subject. Figure 12b shows the corresponding artificial occlusion images of (a).

We carried out the following two experiments in CMU PIE database. In Exp. 1, a random subset with $d(=10,15,20$ and 25) images per individual is taken with labels to form the training set, and the rest of the database is considered to be the testing set. We summarize the best recognition rates of these methods in Table 8. The dimension that results in the best performance for each method is given in parentheses. In Exp. 2, we performed a similar experiment with the Exp. 1, and the only difference is that all test images were artificially occluded by a dinosaur patch at random position. We summarize the best recognition rates of these methods in Table 9. The dimension that results in the best performance for each method is given in parentheses.

As can be seen from Table 8, we can see that our method LAG-LDA achieves the best performance among all methods, irrespective of the variations of training sample size. In all experiments, LAG-LDA has more than 92 % recognition rate. For each method, we calculate the standard deviation for all the recognition rates. As can be seen from Table 8, our method obtains the second smallest standard deviation as compared to the other methods; however, we can see that our method outperforms the first smallest standard deviation method (HE-LDA) in terms of recognition accuracy for approximately up to 84, 85, 86 and 86 % in different training sample sizes (10, 15, 20 and 25). In addition, LAG-LDA outperforms the second best

method (MSR-LDA) in terms of recognition accuracy for approximately up to 6, 4, 5 and 5 % in different training sample sizes (10, 15, 20 and 25). Therefore, we can believe that our algorithm is more robust to pose changes.

As can be seen from Table 9, we can see that the performances of all methods degrade compared with Table 8. However, we can see that our method LAG-LDA still achieves the best performance among all methods, irrespective of the variations of training sample size. In all experiments, LAG-LDA has more than 84 % recognition rate. For each method, we also calculate the standard deviation for all the recognition rates. As can be seen from Table 9, our method does not obtain the smallest standard deviation as compared to the other methods; however, we can see that our method outperforms the first smallest standard deviation method (HE-LDA) in terms of recognition accuracy for approximately up to 76, 79, 79 and 80 % in different training sample sizes (10, 15, 20 and 25). In addition, LAG-LDA outperforms the second best method (GF-LDA) in terms of recognition accuracy for approximately up to 21, 21, 22 and 20 % in different training sample sizes (10, 15, 20 and 25). Therefore, we can believe that our algorithm is more robust to pose changes with different occlusions.

As can be seen from Tables 8 and 9, we can observe that LAG-LDA obtains better results with or without occlusion; however, the other methods are very sensitive to artificial occlusion. Therefore, LAG-LDA is a more stable method.

Table 9 Best recognition rate comparison on Extended Yale B face Database with different methods (dinosaur occlusion)

Method	Exp. 2				
	10	15	20	25	Std.
None-LDA	52.94 % (67)	53.49 % (67)	54.78 % (67)	55.15 % (67)	1.0460
HE-LDA	8.71 % (55)	9.16 % (67)	10.49 % (67)	10.65 % (64)	0.9639
GF-LDA	63.71 % (67)	67.02 % (67)	67.24 % (66)	70.69 % (67)	2.8513
GB-LDA	42.35 % (67)	50.85 % (67)	51.42 % (67)	53.74 % (67)	4.9859
MSR-LDA	40.33 % (67)	37.69 % (67)	37.99 % (67)	31.71 % (63)	3.6747
TT-LDA	23.06 % (67)	25 % (64)	25.17 % (67)	27.21 % (65)	1.6959
WF-LDA	48.65 % (67)	53.29 % (67)	54.12 % (67)	54.78 % (67)	2.7745
LAG-LDA	84.04 % (66)	87.52 % (67)	88.63 % (67)	90.53 % (67)	2.7264

Bold values indicate the best recognition rate

Table 10 The performance of LAG-LDA with different neighbor radius

Database	Neighbor radius (L)	Recognition rate (%)
AR	1	96.29
	2	97.43
	3	96.86
	4	95.43
	5	93.14
Extended Yale B	1	99.45
	2	99.95
	3	99.73
	4	99.62
	5	99.35
CMU PIE	1	90
	2	93.70
	3	94.79
	4	95.17
	5	94.96

Bold values indicate the best recognition rate

3.5 Parameter analysis

In this section, we will discuss the influence of parameter setting of the neighbor radius L . The performances of LAG-LDA with different neighbor radius are evaluated on the face image databases mentioned above. The first session of each individual is used for the training and the second session for tests in the experiment on the AR database. The random 15 images per class of Extended Yale B database are chosen for training, and the rest for testing. In addition, the random 15 images per class are selected for training, and the remaining images for test on the CMU PIE database.

The recognition rates of LAG-LDA with respect to different neighbor radius on the three databases are shown in Table 10. As can be seen from Table 10, LAG-LDA achieves better result when L is set to 2 in AR and

Extended Yale B. It is worth nothing that LAG-LDA achieves better result when L is set to 4 in CMU PIE. Therefore, we generally choose the neighbor radius $L = 2$. For face database with pose changes, we intuitively believe that the larger neighbor radius can deal with larger local image block, and can better adapt to pose changes. In this paper, we choose the neighbor radius $L = 4$ for CMU PIE.

4 Conclusions

In this paper, we propose an image preprocessing method based on LAG. The LAG operator is actually calculated by integrating more directional gradients. Because of considering more directional gradients, LAG captures more edge information for each pixel of an image and finally generates an LAG image, which achieves a more robust image dissimilarity between images. Then, the LAG image is normalized into a normalized feature vector using the “z-score” method. The dimensionality of the normalized feature vector is reduced by linear discriminant analysis to yield a low-dimensional feature vector. Our experiments show that the proposed method is robust to different facial expressions, illumination variations and occlusions changes (including random occlusion position and different occlusion sizes), and achieves better recognition rate.

References

1. Zhao W, Chellappa R, Phillips PJ, Rosenfeld A (2003) Face recognition: a literature survey. *ACM Comput Surv* 34(4):399–485
2. Poh N, Chan CH, Kittler J et al (2010) An evaluation of video-to-video face verification. *IEEE Trans Inf Forensics Secur* 5(4):781–801
3. Turk M, Pentland A (1991) Eigenfaces for recognition. *J Cogn Neurosci* 3(1):71–86

4. Belhumeur PN, Hespanha JP, Kriegman DJ (1997) Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. *IEEE Trans Pattern Anal Mach Intell* 19(7):711–720
5. He X, Yan S, Hu Y, Niyogi P, Zhang HJ (2005) Face recognition using Laplacianfaces. *IEEE Trans Pattern Anal Mach Intell* 27(3):328–340
6. Goudelis G, Zafeiriou S, Pantic M (2012) Subspace learning from image gradient orientations. *IEEE Trans Pattern Anal Mach Intell* 34(12):2454–2466
7. Goudelis G, Zafeiriou S, Tefas A, Pitas I (2007) Class-specific kernel-discriminant analysis for face verification. *IEEE Trans Inf Forensics Secur* 2(3):570–587
8. Yang J, Frangi AF, Yang J, Zhang D, Jin Z (2005) KPCA plus LDA: a complete kernel fisher discriminant framework for feature extraction and recognition. *IEEE Trans Pattern Anal Mach Intell* 27(2):230–244
9. Baudat G, Anouar F (2000) Generalized discriminant analysis using a kernel approach. *Neural Comput* 12:2385–2404
10. Cevikalp H, Neamtu M, Wilkes M (2006) Discriminative common vector method with kernels. *IEEE Trans Neural Netw* 17(6):1550–1565
11. Timo A, Abdenour H, Matti P (2004) Face recognition with Local Binary Patterns. In: *Proceedings of the European Conference Computer Vision*, Springer, Berlin, pp 269–481
12. Ahonen T, Hadid A, Pietikainen M (2006) Face description with local binary patterns: application to face recognition. *IEEE Trans Pattern Anal Mach Intell* 28(12):2037–2041
13. Chengjun L (2004) Gabor-based kernel PCA with fractional power polynomial models for face recognition. *IEEE Trans Pattern Anal Mach Intell* 26(5):572–581
14. Chengjun L, Wechsler H (2002) Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition. *IEEE Trans Image Process* 11(4):467–476
15. Chengjun L, Wechsler H (2003) Independent component analysis of Gabor features for face recognition. *IEEE Trans Neural Netw* 14(4):919–928
16. Jobson D, Rahman Z, Woodell G (1997) A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans Image Process* 6(7):965–976
17. Zhang T, Tang YY, Fang B, Shang Z, Liu X (2009) Face recognition under varying illumination using gradientfaces. *IEEE Trans Image Process Corresp* 18(11):2599–2606
18. Wang H, Li S, Wang Y (2004) Face recognition under varying lighting conditions using self quotient image. In: *Proc. IEEE Int. Conf. Autom. Face Gesture Recognition*, pp 819–824
19. Gross R, Brajovic V (2003) An image preprocessing algorithm for illumination invariant face recognition. In *Proceedings of the AVBPA*, pp 10–18
20. Tan Xiaoyang, Triggs Bill (2010) Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans Image Process* 19(6):1635–1649
21. Wang B, Li W, Yang W, Liao Q (2011) Illumination normalization based on Weber’s law with application to face recognition. *IEEE Signal Process Lett* 18(8):462–465
22. Wright J, Yang A, Ganesh A, Sastry S, Ma Y (2009) Robust face recognition via sparse representation. *IEEE Trans Pattern Anal Mach Intell* 31(2):210–227
23. Yang M, Zhang L, Yang J, Zhang D (2011) Robust sparse coding for face recognition. In *Proc. IEEE Int’l Conf. Computer Vision and Pattern Recognition*
24. Elhamifar E, Vidal R (2011) Robust classification using structured sparse representation. In: *Proceedings of the IEEE Int’l Conf. Computer Vision and Pattern Recognition*
25. Yang M, Zhang L, Yang J, Zhang D (2013) Regularized robust coding for face recognition. *IEEE Trans Image Process* 22(5):1753–1766
26. Martinez AM, Benavente R (1998) The AR face database. CVC technical report
27. Lee KC, Ho J, Kriegman DJ (2005) Acquiring linear subspaces for face recognition under variable lighting. *IEEE Trans Pattern Anal Mach Intell* 27(5):684–698
28. Sim T, Baker S, Bsat M (2003) The CMU pose, illumination, and expression database[J]. *IEEE Trans Pattern Anal Mach Intell* 25(12):1615–1618
29. Guan N, Tao D, Luo Z et al (2012) MahNMF: Manhattan non-negative matrix factorization[J]. *arXiv preprint arXiv:1207.3438*
30. Zafeiriou S, Tefas A, Buciu I, Pitas I (2006) Exploiting discriminant information in nonnegative matrix factorization with application to frontal face verification. *IEEE Trans Neural Networks* 17(3):683–695
31. Guan Naiyang, Tao Dacheng, Luo Zhigang, Yuan Bo (2011) Manifold regularized discriminative non-negative matrix factorization with fast gradient descent. *IEEE Trans Image Process* 20(7):2030–2048
32. Jain A, Nandakumar K, Ross A (2005) Score normalization in multimodal biometric systems. *Pattern Recogn* 38(12):2270–2285
33. Jun Yu, Yong Rui, Yuan Yan Tang, Dacheng Tao (2014) High-order distance-based multiview stochastic learning in image classification. *IEEE Trans Cybern* 44(12):2431–2442
34. Jun Yu, Tao Dapeng, Li Jonathan, Cheng Jun (2014) Semantic preserving distance metric learning and applications. *Inf Sci* 281(10):674–686