Local Structure-Based Image Decomposition for Feature Extraction With Applications to Face Recognition

Jianjun Qian, Jian Yang, *Member, IEEE*, and Yong Xu, *Member, IEEE*

***Abstract*— This paper presents a robust but simple image feature extraction method, called image decomposition based on local structure (IDLS). It is assumed that in the local window of an image, the macro-pixel (patch) of the central pixel, and those of its neighbors, are locally linear. IDLS captures the local structural information by describing the relationship between the central macro-pixel and its neighbors. This relationship is represented with the linear representation coefficients determined using ridge regression. One image is actually decomposed into a series of sub-images (also called structure images) according to a local structure feature vector. All the structure images, after being down-sampled for dimensionality reduction, are concatenated into one super-vector. Fisher linear discriminant analysis is then used to provide a low-dimensional, compact, and discriminative representation for each super-vector. The proposed method is applied to face recognition and examined using our real-world face image database, NUST-RWFR, and five popular, publicly available, benchmark face image databases (AR, Extended Yale B, PIE, FERET, and LFW). Experimental results show the performance advantages of IDLS over state-of-the-art algorithms.**

***Index Terms*— Image decomposition, local structure feature, ridge regression, face recognition.**

I. Introduction

FACE recognition is a popular research field in pattern recognition, and has attracted much attention in many areas, such as information security, law enforcement, human–computer interaction and surveillance [1]. In real-world applications, it's difficult to deal with the range of appearance variations that commonly occur in face images owing to pose, illumination and facial expression. With these complexities in face representation, the main challenge is how to exploit robust and discriminative features. A large number of feature extraction algorithms have been developed and tested for face recognition, including subspace based global features and local appearance features. Among them, principal component analysis (PCA) [30] and Fisher linear discriminant analysis (FLDA) [2] — two well-known linear subspace learning methods which have been widely used in the pattern recognition and computer vision areas — have become the most popular techniques for face recognition. Simultaneously, the kernel methods (e.g. kernel principal component analysis [3] and kernel Fisher discriminant [4] etc.) are introduced to deal with the nonlinear structure of data. These kernel methods have proven to be effective in many real-world applications [5], [6]. Recently, Yan *et al.* [7] proposed a general framework called graph embedding. Various manifold-based learning methods such as ISOMAP, LLE, Laplacian Eigenmap, and LPP etc [8]–[12], can be reformulated as a unified model in this framework. Some experiments have shown that these methods can find perceptually meaningful embeddings for face images.

Compared with the global features like PCA and FLDA, local appearance features are more robust for handling local changes such as illumination, expression and pose. Gabor wavelets [13] and local binary pattern [14] are two popular methods for feature extraction and image representation in the last decade. The Gabor wavelet has been extensively used in face recognition due to the fact that its kernels are similar to the two-dimensional receptive field profiles of the mammalian cortical simple cells. Also, it captures the desirable local characteristic structure of spatial frequency, spatial locality, and selective orientation, which are proven to be robust to illumination and facial expression changes. Local Binary Pattern (LBP) has received increasing attention owing to its simplicity and thus far encouraging performance in texture image-analysis tasks [14]. Many approaches derived from LBP have been proposed and successfully used in various applications [15], [16]. In particular, Ahonen *et al.* proposed a method based on local binary pattern histograms, which has been demonstrated effective for face recognition. The representation of a face image is obtained by dividing the image into a series of sub-windows and computing their histograms of LBP values [16], [17]. Subsequently, some works combined LBP and Gabor features to improve the face recognition performance compared to the individual use

Manuscript received July 15, 2011; revised November 11, 2011, March 27, 2012, July 16, 2012, January 9, 2013, and April 10, 2013; accepted May 8, 2013. Date of publication May 22, 2013; date of current version August 9, 2013. This work was supported in part by the National Science Fund for Distinguished Young Scholars under Grant 61125305 and Grant 61233011, the Key Project of Chinese Ministry of Education. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Bulent Sankur.

J. Qian is with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: qjjtx@126.com).

J. Yang is with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China, and also with Computation and Neural Systems, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: jianyang@caltech.edu).

Y. Xu is with the Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen 518055, China (e-mail: laterfall2@yahoo.com.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2013.2264676

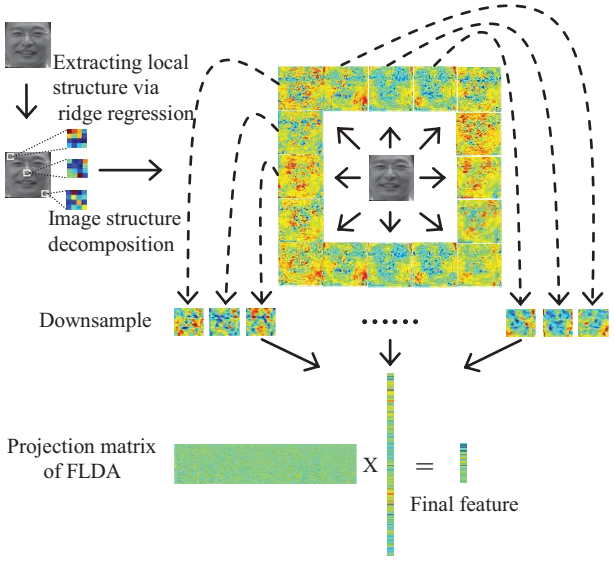


Fig. 1. The overview of the IDLS method.

of any one of these representations. Zhang *et al.* proposed the local Gabor binary pattern method and used a histogram sequence to model facial images [18]. Zhao and Pietikäinen presented a method called LBP-TOP for dynamic texture and facial analysis [19]. Lei *et al.* developed a face-representation method which not only exploits the information in the spatial domain but also takes into consideration different scales and orientations [20]. There are also other approaches [21], [22] using Gabor phase information. Moreover, a family of novel face image descriptors (e.g. three-patch LBP (TPLBP) and four-patch LBP (FPLBP)) has been developed, and they can capture certain statistics of local patch similarities [31], [32].

By far, numerous methods have been developed for image decomposition, such as morphological component analysis (MCA), sparse representation and structure-texture decomposition model. These methods can be applied to medical imaging, image compression, image restoration and image analysis [23]–[26]. As we know, most images contain rich local structural information, which can potentially be used in object recognition tasks. Therefore, we consider decomposing an image into a series of sub-images which contain different structural information. Based on this idea, we propose a simple but effective feature extraction method called image decomposition based on local structure (IDLS). IDLS first captures the local structural information by characterizing the relationship between the macro-pixel of a central pixel and those of its neighbors (one neighbor in one orientation). Here, a linear regression model is used to assess relationships among them. After that, we obtain an N-dimensional feature vector for each pixel. One image is therefore decomposed into N sub-images (also called structure images). Then, we down-sample all structure images and normalize each to a unit vector. The normalized structure images are concatenated into one super-vector. The dimensionality of the super-vector is further reduced by FLDA to yield a compact, low-dimensional feature vector. The overview of the proposed method is shown in Fig. 1.

It is necessary to highlight two merits of the proposed method. First, IDLS uses a set of linear representation coefficients to characterize the relationships between a central pixel and its neighbors. This set of linear representation coefficients is determined by optimizing a linear regression model. Second, IDLS is robust to illumination changes. This is because IDLS computes the local structure relationships between a central pixel and its neighbors in a local window. Any intensity change of all the pixels of an image will not affect these local structure relationships.

The rest of the paper is organized as follows. In Section II, we present the details of the proposed IDLS method and analyze the advantage of IDLS feature over Gabor feature. We experiment using our real-world face image database NUST-RWFR and five popular publicity face image databases in Section III. The experimental results demonstrate that the effectiveness of the proposed method. Finally, a conclusion is offered in Section IV.

II. Image Decomposition Based on Local Structure

A. Outline of Ridge Regression

Linear regression attempts to model the relationship between the dependent variables and the explanatory variables (regressors) by fitting a set of linear equations to observed data. Specifically, given a dataset $\{y_i, x_{i,1}, \cdots, x_{i,P}\}_{i=1}^{n}$, the linear regression model takes the following form:

$$y_i = w_i x_{i,1} + \cdots + w_P x_{i,P} + \varepsilon_i, \qquad i = 1, \cdots, n, \tag{1}$$

where y_i is the dependent variable, $x_{i,1}, \cdots, x_{i,P}$ are explanatory variables, and ε_i is the disturbance term.

Actually, these n equations can be stacked together and written in vector form as:

$$\mathbf{y} = \mathbf{X}\mathbf{w} + \varepsilon, \tag{2}$$

where $\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, $\mathbf{X} = \begin{pmatrix} x_{1,1} & \cdots & x_{1,P} \\ x_{2,1} & \cdots & x_{2,P} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,P} \end{pmatrix}$, $\mathbf{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_P \end{pmatrix}$, and $\varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$.

Generally, the parameter $\mathbf{w}$ can be determined by solving the following optimization problem:

$$\min_{w} \|\mathbf{y} - \mathbf{X}\mathbf{w}\|^2 . \tag{3}$$

The ordinary least squares solution of Eq. (3) is given by

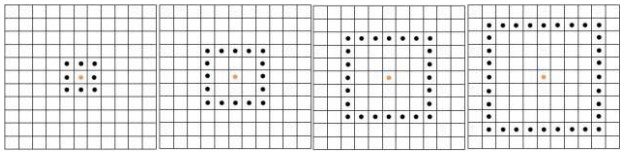
$$\hat{w} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}. \tag{4}$$

However, $\mathbf{X}^T\mathbf{X}$ is singular when the system is underdetermined (i.e. n<P). One common way to address this problem is to solve the following regularized linear regression model instead

$$\min_{w} \|\mathbf{y}-\mathbf{X}\mathbf{w}\|^2 + \delta \|\mathbf{w}\|^2 . \tag{5}$$

The model is generally called Ridge Regression. Its solution is given by

$$\hat{\mathbf{W}} = (\mathbf{X}^T\mathbf{X} + \delta\mathbf{I})^{-1}\mathbf{X}^T\mathbf{y}. \tag{6}$$



(a) P=8 (R=1) (b) P=16 (R=2) (c) P=24 (R=3) (d) P=32 (R=4)

Fig. 2. Squarely symmetric neighbor sets for radius R. (a) $P = 8$ (R = 1), (b) $P = 16$ (R = 2), (c) $P = 24$ (R = 3), and (d) $P = 32$ (R = 4).

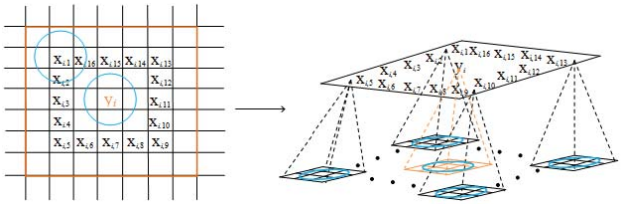


Fig. 3. Illustration of local macro-pixels in a local window

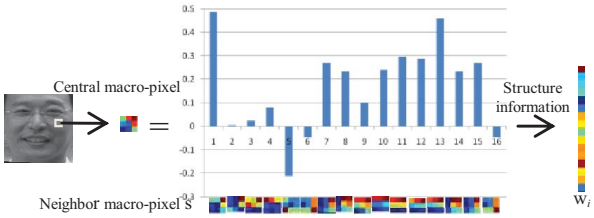


Fig. 4. The procedure of extracting local structure information. Histogram demonstrates the relationship between the central macro-pixel (the macro-pixel of i-th pixel) and its neighbor macro-pixels (Note $\mathbf{w}_i$ is the local structure feature vector of i-th pixel).

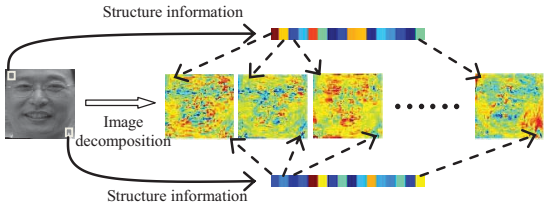


Fig. 5. Image decomposition according to local structure information.

B. Extracting Local Structure Feature Using Ridge Regression

To describe the local structure, let us begin with the characterization of neighbor sets, local macro-pixels, and local windows. Suppose there are N pixels in an image. We treat the i-th pixel in an image as a center and determine its P neighbor pixels on a square of radius R using the city-block distance[1]. These neighbor pixels form a squarely symmetric neighbor set $\Omega_P^i = \{i_j | j = 1, \cdots, P\}$. Fig. 2 illustrates these neighbor sets for different R.

We treat the i_j-th pixel belonging to the neighbor set Ω_P^i as a center to choose a $K \times K$ local macro-pixel (e.g. K=3, 5). The gray values corresponding to all K^2 pixels within the macro-pixel form a Q-dimensional local macro-pixel vector $\mathbf{x}_{i,j}$, $j = 1, \cdots, P$, and $Q = K^2$. Similarly, for the center pixel of the neighbor set Ω_P^i, we can choose a same-sized, central macro-pixel which is denoted by a Q-dimensional vector $\mathbf{y}_i$. For the i-th pixel in the image, all of the chosen local macro-pixels, $\mathbf{y}_i$ and $\mathbf{x}_{i,j}$ $(i = 1, \cdots, P)$, determine a local window of the pixel. Fig. 3 shows an example of a local window which contains a central macro-pixel and 16 neighbor macro-pixels. Notice that the size of neighbor set P is 16 and the size of macro-pixel Q is $3 \times 3 = 9$.

Here we assume that all macro-pixels in a local window are locally linear[2]. Based on this assumption, we characterize the local structure between a central macro-pixel and its neighbor macro-pixels. Specifically, given a central macro-pixel and its neighbor macro-pixels in the i-th pixel centralized local window, the central macro-pixel $\mathbf{y}_i$ can be approximately represented by a linear combination of neighbor macro-pixels $\mathbf{x}_{i,1}, , \cdots \mathbf{x}_{i,P}$ as follows

$$\mathbf{y}_i = w_{i,1}\mathbf{x}_{i,1} + \cdots + w_{i,P}\mathbf{x}_{i,P} + \varepsilon, \tag{7}$$

where $w_{i,1}, \cdots, w_{i,P}$ are the representation coefficients.

Let $\mathbf{w}_i = [w_{i,1}, \cdots w_{i,P}]^T$ and $\mathbf{X}_i = [\mathbf{x}_{i,1}, , \cdots \mathbf{x}_{i,P}]$, then the Eq. (7) can be rewritten as

$$\mathbf{y}_i = \mathbf{X}_i\mathbf{w}_i + \varepsilon. \tag{8}$$

To avoid the singularity problem, we use the ridge regression to get a solution of Eq. (8), i.e.

$$\mathbf{w}_i = (\mathbf{X}_i^T\mathbf{X}_i + \delta\mathbf{I})^{-1}\mathbf{X}_i^T\mathbf{y}_i. \tag{9}$$

Actually, the vector $\mathbf{w}_i$ describes the structural relationship of the i-th pixel and its neighbor pixels in the current local window. Fig. 4 illustrates the procedure of extracting local structural information of one pixel from an image.

C. Image Decomposition and Feature Representation

From section II.B, for each pixel in an image, we can obtain a local structure feature vector to describe the local structural relationship between the central macro-pixel and its neighbor macro-pixels. Assuming that there are N pixels in an image, the local structure feature vectors of all pixels, $\mathbf{w}_1, \cdots, \mathbf{w}_N$, form a local structure feature matrix $\mathbf{W}$ as follows

$$\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_N] = \begin{pmatrix} w_{1,1} & w_{2,1} & \cdots & w_{N,1} \\ \vdots & \vdots & \ddots & \vdots \\ w_{1,P} & w_{2,P} & \cdots & w_{N,P} \end{pmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \vdots \\ \mathbf{v}_P \end{bmatrix} \tag{10}$$

Each row of $\mathbf{W}$, an N-dimensional vector $\mathbf{v}_j$, can be reformed into an image, which is called structure image. Thus, the local structure feature matrix yields a set of P structure images. In other words, one image is decomposed into P structure images by virtue of the local structural information. Fig. 5 shows structure images derived from one given image. Since each of structure images is composed of the structural information between the same orientation neighbors and their central pixels, they represent structural feature from different orientations effectively. From Fig. 3, we see that the current central pixel has one neighbor in one orientation in the local

[1]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.

[2]Actually, a similar assumption was used in LLE [9].

window. Each structure image accumulates almost all the structural information in one orientation by this means.

All the structure images are concatenated into one augmented super-vector so as to include all structural information from different orientations. Before the concatenation, each structure image is down-sampled by a factor λ to reduce feature dimension and then normalized to be unit vector. Letting $\mathbf{v}_p^\lambda$ $(p = 1, \cdots, P)$ represent the p-th down-sampled and normalized structure image vector, the augmented super-vector $\mathbf{V}^\lambda$ is thus defined as follows:

$$\mathbf{V}^\lambda = (\mathbf{v}_1^\lambda, \mathbf{v}_2^\lambda, \cdots, \mathbf{v}_P^\lambda). \tag{11}$$

Notice that the augmented super-vector $\mathbf{V}^\lambda$ is still high dimensional. To deal with the high-dimensional problems, numerous dimensionality reduction methods have been developed in recent years. The basic idea of these methods is to seek a meaningful low dimensional subspace in a higher input data space. This subspace can provide a compact representation of higher dimensional data. Here, FLDA[3] is used to further reduce the dimension of the augmented super-vector and to generate compact and discriminative feature vectors. From this point of view, our method not only considers the local structure, but also takes into account the global information.

D. IDLS via Gaussian Kernel Distance

We know that LLE [9] compute the local weights using the linear regression, while a Laplacian Eigenmap [10] and LPP [11] calculate the local weights directly using the Gaussian kernel distance. In contrast to IDLS, which characterizes the local structure relationship via linear regression coefficients, we here present a way to characterize the local structure relationship by computing the distances between the central macro-pixel and each neighbor macro-pixel with a Gaussian kernel. Specifically, in the i-th pixel centralized local window, the distance between the central macro-pixel $\mathbf{y}_i$ and the j-th neighbor macro-pixel $\mathbf{x}_{i,j}$ $(j = 1, \cdots, P)$ is defined as follows

$$w_{i,j} = \exp\{-\frac{||\mathbf{y}_i - \mathbf{x}_{i,j}||^2}{h}\} \tag{12}$$

where h is the smoothing parameter (h is set to 0.25 in this paper). The distance $w_{i,j}$ is used to describe the structural relationship between the central pixel and the j-th neighbor. Therefore $\mathbf{w}_i = [w_{i,1}, \cdots, w_{i,P}]^T$ represent the structural information of the i-th pixel in the local window. Based on the structural information of all pixels in an image, we can obtain a set of P structure images. This method is called IDLS-Distance, since it uses the Gaussian kernel distance to characterize the local structure relationship.

Note that IDLS-Distance considers the relationship between the central macro-pixel and each neighbor macro-pixel in isolation. We know that IDLS exploits the structural information using linear regression coefficients as shown in Eq. (9). That is, IDLS takes into account the relationship between central macro-pixel and all neighbor macro-pixels together. In other words, IDLS takes advantage of the information of all neighbor macro-pixels to help balance the relationship between the central macro-pixel and each neighbor macro-pixel.

[3]Note that FLDA is degenerated into PCA when there is only one image per person for training.

E. Comparisons With Gabor Feature

IDLS and Gabor both decompose one image into a series of structure images according to local structural information. All the structure images, after being down-sampled for dimensionality reduction, are concatenated into one super-vector. FLDA is then used to obtain the low-dimensional, compact feature vector. Gabor feature representation of an image is made via convolution operation with family kernels, which stimulate the 2-D receptive filed profiles of mammalian cortile simple cell. Moreover, they represent desirable characteristics of spatial scale, spatial locality and selective orientation. Specifically, one would use Gabor kernels of five different scales and eight orientations to generate forty sub-images in general. Every structure image is down-sampled to reduce space dimensions. Afterwards, we concatenate all the represented results and develop an augmented super-vector to encompass significant features of different scales, localities and orientations. IDLS, however, first exploits the local structure hidden in the unobserved space of an image by using the linear regression coefficients to depict the relationships between the macro-pixels of each pixel and those of their neighbors. From the reconstruction point of view, it's believed that linear regression coefficients reveal intrinsic local structure relationship between a central macro-pixel and its neighbors. One image is actually decomposed into a series of structure images. This is the difference between IDLS and Gabor feature.

To show the advantage of IDLS compared with Gabor feature under varying lighting conditions, we choose three macro-pixels (5×5) with the same position in three images under different lighting conditions, as shown in Fig. 6. Then, structural information is achieved by decomposing the image via ridge regression and Gabor filtering, respectively. By this means, the difference of structural information of the macro-pixels with different illuminations is calculated as shown in Fig. 6. From Fig. 6, we can see that structural information of a Gabor feature is sensitive to the lighting changes. On the other hand, our method using ridge regression to represent structural information is robust to the lighting changes. Fig. 7 shows the Gabor magnitude face images and structure images of IDLS with respect to the image c in Fig. 6. From Fig. 7, we can see that IDLS has a power to eliminate the effect of illumination. Because IDLS captures the local structural information of each pixel in its local window, and an intensity change is added to each image pixel will have little effect on local structural information. Therefore, IDLS demonstrates its advantages over Gabor under different lighting conditions. Experiments in Section III will demonstrate our ideas mentioned above.

III. Experiments

In this section, we will evaluate the effectiveness of IDLS and compare it with some state-of-the-art algorithms by experimenting on our real world face image database NUST-RWFR and five large publicly available face image databases

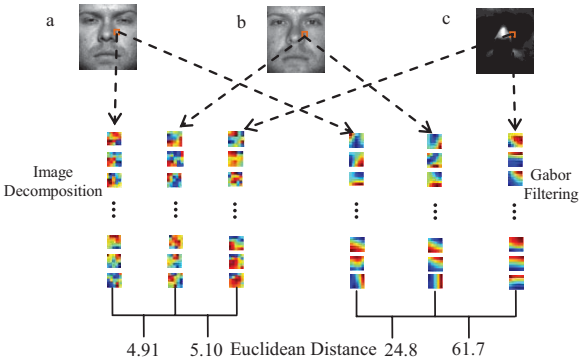


Fig. 6. The differences of the structure information of the macro-pixels with different illuminations.

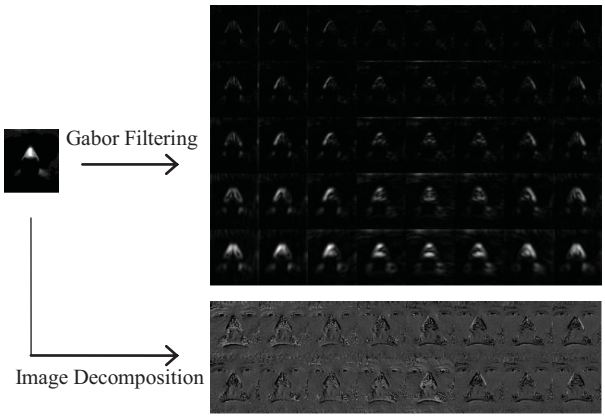


Fig. 7. The Gabor magnitude face images and structure images of IDLS.

(AR, Extended Yale B, PIE and FERET). In the following experiments, we choose the regularized parameter in ridge regression as $\delta = 0.01$,[4] and the down- sampling factor as $\lambda = 2$ in our IDLS method. In addition, the optimal parameters of the state-of-the-art algorithms in our experiments are listed as follows. For the LBP features[5] [17], the sub-block size is 8×8 on the NUST-RWFR, Extended Yale B and PIE, 16×15 on the AR. Each block is with a 59-D histogram. The parameters of three patch based LBP (TPLBP) [32] descriptors are $r_1 = 2$, S = 8, and $w = 5$. The MATLAB source code is from the Web.[6] For the LTP and LTP/DT features[7] [15], the threshold for coding is 0.02, the α parameter for DT distance is 1, the threshold for truncating DT distance is 6, and the sub-block size is 16×15 on the AR dataset and 8×8 on the other datasets. For ID-LARK [36] and LARK[8] [37], the local window size is 7×7 in all experiments except for the one on the Extended Yale B dataset where the local window size is 3×3. To compute the Gabor feature [13], we used 40 Gabor filters including 5 scales and 8 directions for each face image. The down-sampling factor is 4 on the NUST-RWFR,

[4]Actually, we can obtain the similar results when $\delta \in [0.001, 0.01]$ in our experiments.

[5]http://www.ee.oulu.fi/mvg/page/lbpmatlab.

[6]http://www.openu.ac.il/home/hassner/projects/Patchlbp/.

[7]http://parnec.nuaa.edu.cn/xtan/Publication.htm.

[8]http://users.soe.ucsc.edu/~milanfar/research/rokaf/.html/download_object.php.

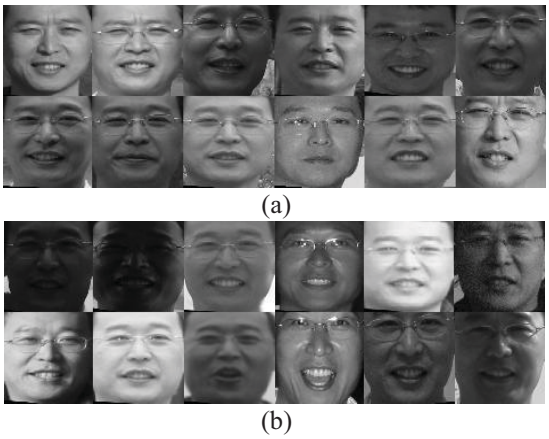


Fig. 8. Sample images for one subject of NUST-RWFR database (a) session 1, (b) session 2.

PIE, FERET, LFW datasets and 8 on the AR, Extended Yale B datasets.

A. Experiment Using the NUST-RWFR Database

The NUST-RWFR database[9] contains 2400 color face images of 100 persons, including frontal views of faces with different facial expressions, lighting conditions and degrees of blurring. All the pictures are taken in a real world situation. The pictures of 100 persons were taken in two sessions and each session contains 12 color images. The quality of pictures in the first session is good, and that in the second session is poor. All face images of these 100 persons are used in our experiments. The face portion of each image is manually cropped and then normalized to 80×80 pixels. The sample images of one person as shown in Fig. 8. Additionally, we resize the face images into 24×24 and 48×48 for further experiments.

In the first experiment, the images from the first session of each person are used for training, and images from the second session for testing. The total training samples is 1200, since there are 12 face images in session 1 of each person. PCA, FLDA, LARK combines binary-like representation and matrix cosine similarity (MCS) [37], ID-LARK, LBP plus Chi2, LTP plus Chi2, LTP/DT, TPLBP plus Chi2, Gabor plus FLDA, and proposed methods IDLS-Distance, IDLS are used for image feature extraction, respectively. We thus use PCA to reduce the dimension of each image feature vector to be 200 before implementing FLDA. Finally, the nearest neighbor (NN) classifier is used for classification.

Table I lists the recognition rates of each method and the entire running time under three different image sizes. From Table I, we can see that the proposed method IDLS is significantly better than others. Simultaneously, IDLS-Distance outperforms PCA, FLDA, LARK, ID-LARK, LBP plus Chi2, LTP plus Chi2, LTP/DT and TPLBP plus Chi2. It is comparable with Gabor plus FLDA in general. So, IDLS-Distance still reveals that image decomposition which uses local structural information can exploit the structure

[9]http://pcalab.njust.edu.cn/.

TABLE I

THE RECOGNITION RATE (%) OF PCA, FLDA, LARK, ID-LARK, LBP+CHI2, LTP+CHI2, LTP/DT, TPLBP+CHI2, GABOR+FLDA, IDSL-DISTANCE AND IDLS WITH THE NN CLASSIFIER, AND THE CPU TIME FOR RUNNING ON THE NUST-RWFR DATABASE USING DIFFERENT IMAGE SIZES

Methods	24 × 24		48 × 48		80 × 80	
	Recognition Rate	Time(s)	Recognition Rate	Time(s)	Recognition Rate	Time(s)
PCA	16.9	24	17.0	36	*17.7*	*48*
FLDA	45.8	34	45.3	34	*45.0*	*35*
LARK	42.2	269	45.8	988	*47.2*	*3158*
ID-LARK	50.6	236	55.7	475	*56.2*	*1011*
LBP+Chi2	32.8	97	48.1	265	*49.2*	*286*
LTP+Chi2	40.9	92	45.4	203	*43.8*	*335*
LTP/DT	43.1	4137	47.1	6582	*48.2*	*9823*
TPLBP+Chi2	20.4	98	33.4	235	*36.9*	*607*
Gabor+FLDA	55.9	187	65.8	389	*66.7*	*617*
IDLS-Distance	60.7	165	*64.7*	*492*	65.3	1407
IDLS	67.7	202	*73.8*	*601*	75.0	1569

feature of an image efficiently. The structure feature helps to improve the performance of the recognition task. Moreover, the structure layout (there is one neighbor in one direction of the central macro-pixel in the local window) ensures that structural information would not be weakened by different structure images from the same orientation. Compared with IDLS-Distance, IDLS assumes that central macro-pixel and its neighbor macro-pixels in a local window are locally linear. We thus use neighbor macro-pixels to represent the central macro-pixel and the ridge regression is applied to solve the representation coefficients. These coefficients, which describe the relationship between central macro-pixel and its neighbors, are considered as structural information in the local window. By this means, IDLS considers synthetically the relationship during the macro-pixels in the local window. Nevertheless, IDLS-Distance only computes the distance between central macro-pixel and one neighbor macro-pixel with Gaussian kernel as structural information. Therefore, this structural information is weaker than that of IDLS. This is the reason IDLS gives better performance than IDLS-Distance. In addition, Gabor is widely used owing to the fact that it describes the local feature by using family kernels (forty kernels with five different scales and eight different orientations) to convolve with the image. Despite this, the performance of IDLS is about 8% higher than that of Gabor plus FLDA. These demonstrate that our method using linear regression coefficients can describe the local structure well and improve the recognition rate.

In terms of CPU time for running, IDLS consumes more time than the other methods, since it calculates a least square solution once in the process of computing the local structural information of each pixel in an image. However, we can significantly improve the computational efficiency of IDLS by resizing the image into a small size before implementing it, noticing that IDLS can still give better results using the low-resolution face images than the state-of-the-art methods using the high-resolution images. Fig. 9 illustrates that IDLS, using the 48 × 48 images, consistently performs better than FLDA and Gabor plus FLDA using the 80 × 80 images. Actually, a similar result can be seen in Table I. From this

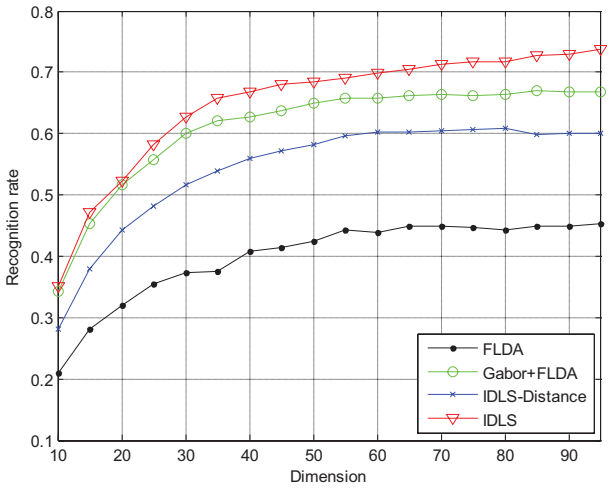


Fig. 9. The recognition rates of FLDA, Gabor+FLDA, IDLS-Distance and IDLS with the variations of dimensions on the NUST-RWFR database. (The first session of each subject forms the training set, and the second session of each subject forms the testing set).

point of view, it's necessary to point out that face images are all down-sampled into a smaller image size of 48 × 48 before implementing the proposed method in the following experiments on the other face image databases.

Furthermore, we randomly choose K (K varies from 4 to 12 with interval 2) images from each class for training; the rest of the images are used for testing. We perform 10 runs of tests with different face image sizes for each K. The average recognition rates of PCA, FLDA, LARK, ID-LARK, LBP plus Chi2, LTP plus Chi2, LTP/DT, TPLBP plus Chi2, Gabor plus FLDA, IDLS-Distance and IDLS with the NN classifier are shown in Fig. 10. The results in Fig. 10 are generally consistent with those in Table I. IDLS always shows better results than the other methods, but IDLS-Distance is not as good as that of Gabor plus FLDA in some cases. This exposes the shortcoming of the IDLS-Distance to a certain degree. It is not very robust to use Gaussian weights between the central macro-pixel and its neighbor macro-pixels to depict

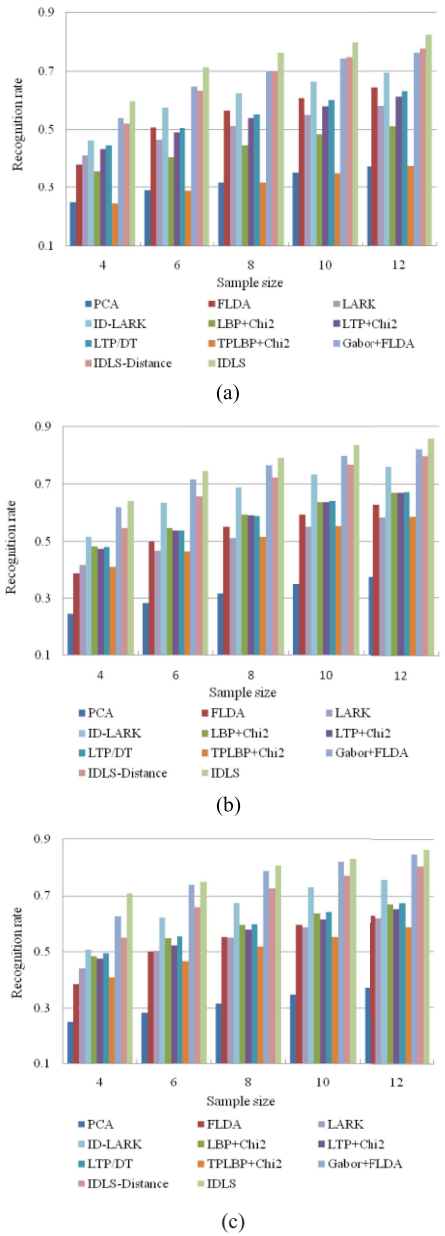


Fig. 10. The average recognition rates of PCA, FLDA, LARK, ID-LARK, LBP plus Chi2, LTP plus Chi2, LTP/DT, TPLBP plus Chi2, Gabor plus FLDA, IDLS-Distance and IDLS with the NN classifier. (a) On the 24 × 24 image set; (b) on the 48 × 48 image set; (c) on the 80 × 80 image set.

the local structural information. In contrast, IDLS is efficient and robust to utilize linear regression coefficients for local structure representation.

B. Experiment Using the AR Database

The AR face database [33] contains over 4000 color face images of 126 persons (70 men and 56 women), including frontal views of faces with different facial expression, lighting conditions and occlusions.[10] The pictures of 120 individuals (65 men and 55 women) were taken in two sessions (separated by two weeks) and each session contains 13 color images.

[10]http://www2.ece.ohio-state.edu/~aleix/ARdatabase.html.

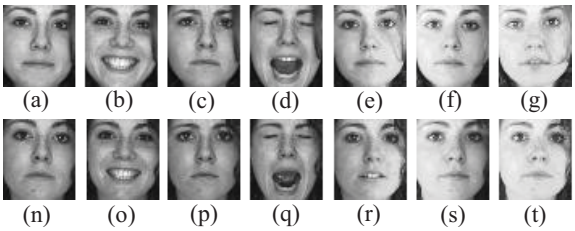


Fig. 11. Sample images for one person of AR database.

TABLE II
THE RECOGNITION RATE OF PCA, FLDA, LARK, ID-LARK, LBP+CHI2, LTP+CHI2, LTP/DT, TPLBP+CHI2, GABOR+FLDA, IDLS-DISTANCE, AND IDLS WITH THE NN CLASSIFIER ON THE AR DATABASE

Methods	Recognition Rate
PCA	60.7
FLDA	78.1
LARK	82.5
ID-LARK	91.9
LBP+Chi2	91.3
LTP+Chi2	90.9
LTP/DT	89.5
TPLBP+Chi2	91.6
Gabor+FLDA	91.4
IDLS-Distance	93.5
IDLS	97.4

Fourteen face images (each session contains 7) of these 120 individuals are selected and used in our experiment. The face portion of each image is manually cropped and then normalized to 100 × 90 pixels. The sample images of one person as shown in Fig. 11, where (a)–(g) are from session 1, and (n)–(t) are from session 2. The details of images are: (a) neutral expression, (b) smiling, (c) angry, (d) screaming, (e) left light on, (f) right light on, (g) all sides light on; and (n)–(t) were taken under the same conditions as (a)–(g).

In our experiment, images from the first session (i.e., (a)–(g)) are used for training, and images from the second session (i.e., (n)–(t)) are used for testing. Then, PCA, FLDA, LARK, ID-LARK, LBP plus Chi2, LTP plus Chi2, LTP/DT, TPLBP plus Chi2, Gabor plus FLDA, IDLS-Distance and the proposed method IDLS are employed for feature extraction, respectively. In the PCA phase of FLDA, the number of principal component is set 200. Further, the NN classifier is used for classification. The recognition results of all methods are given in Table II. These results are almost consistent with the experimental results in Sections III.A. The proposed IDLS performs better than other state-of-the-art methods. The performances of PCA, FLDA, and LARK are not satisfying. It appears that these three methods are sensitive to face image variations over time. In contrast, Gabor feature, ID-LARK, LBP, TPLBP, LTP and IDLS are more robust to time changes and thus achieve better results.

C. Experiment Using the Extended Yale B Database

The extended Yale B face image database [27] contains 38 human subjects under 9 poses and 64 illumination

TABLE III

THE AVERAGE RECOGNITION RATE (PERCENT) AND STD OF PCA, FLDA, LARK, ID-LARK, LBP+CHI2, LTP+CHI2, LTP/DT, GABOR+FLDA, IDLS-DISTANCE, AND IDLS WITH THE NN CLASSIFIER ON THE EXTENDED YALE B DATABASE

Methods	4	8	12	16	20	24
PCA	41.6 ± 5.77	59.7 ± 6.26	65.5 ± 7.11	70.6 ± 2.86	75.2 ± 3.09	78.2 ± 4.92
FLDA	57.3 ± 11.8	77.5 ± 7.91	86.7 ± 4.10	91.1 ± 2.86	92.3 ± 3.21	94.1 ± 3.38
LARK	84.5 ± 5.30	90.6 ± 2.20	92.2 ± 2.83	92.6 ± 3.20	93.7 ± 3.06	94.9 ± 2.35
ID-LARK	86.8 ± 4.73	91.3 ± 3.43	96.4 ± 1.63	96.9 ± 0.88	97.7 ± 0.54	97.9 ± 0.47
LBP+Chi2	84.2 ± 8.59	93.2 ± 1.68	95.3 ± 1.71	96.3 ± 1.02	96.6 ± 0.91	96.7 ± 1.26
LTP+Chi2	62.1 ± 8.79	76.7 ± 5.08	81.9 ± 3.76	84.5 ± 3.75	85.3 ± 2.52	87.1 ± 3.52
LTP/DT	64.3 ± 10.8	80.8 ± 5.08	86.6 ± 3.73	89.6 ± 3.55	91.1 ± 2.14	92.3 ± 2.61
TPLBP+Chi2	50.3 ± 11.7	66.3 ± 5.67	71.8 ± 5.42	75.4 ± 2.81	76.5 ± 3.36	78.9 ± 5.14
Gabor+FLDA	60.3 ± 3.94	78.2 ± 3.98	85.2 ± 3.10	91.2 ± 3.48	93.3 ± 2.83	94.3 ± 1.65
IDLS-Distance	74.9 ± 7.49	84.4 ± 4.16	91.3 ± 2.10	93.5 ± 2.01	93.9 ± 1.69	95.1 ± 1.65
IDLS	95.8 ± 1.73	96.9 ± 0.72	97.9 ± 0.50	98.0 ± 0.66	98.0 ± 0.75	98.2 ± 0.84

Fig. 12. Samples of a person under different illuminations in the Extended Yale B face database.

conditions. The 64 images of a subject in a particular pose are acquired at camera frame rate of 30 frames/second, so there is only small change in head pose and facial expression for those 64 images. However, its extreme lighting conditions still make it a challenging task for most face recognition methods. All frontal-face images marked with P00 are used in our experiment, and each is resized to 96 × 84 pixels. Some sample images of one person are shown in Fig. 12.

In the first experiment, we randomly choose K samples from each class for training and the remaining images for test. Here, K varies from 4 to 24 with interval of 4. For each K, we perform 10 runs of tests for each method mentioned above. Before implementing FLDA, we use PCA to reduce the dimension to be 200 with respect to different number of training samples per class. Also, the NN classifier is employed for classification. The average recognition rates and *std* of each method across ten tests are shown in Table III. Fig. 13 shows the recognition rate curve versus the variation of training sample size.

From Fig. 13, we can see that the recognition rates of PCA are relatively poor. Although TPLBP can give better performance in face verification, experimental result shows that it is not robust for face recognition under different lighting conditions. The performances of Gabor plus FLDA and LARK are not very satisfying. LBP descriptor achieves better results than the above methods. IDLS still outperforms all subspace learning methods and local appearance feature extraction algorithms. This demonstrates that IDLS is the most robust method with respect to illumination changes. It is interesting that here ID-LARK achieves very good results, particularly when the training sample size becomes larger. ID-LARK also uses the local structural information to decompose the image. It captures the local structural information by measuring the geodesic distances between the central

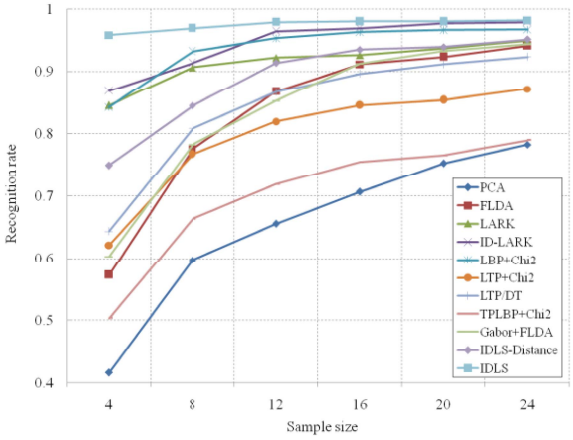


Fig. 13. The average recognition rates of PCA, FLDA, LARK, ID-LARK, LBP+Chi2, LTP+Chi2, LTP/DT, TPLBP+Chi2, Gabor+FLDA, IDLS-Distance and IDLS versus the variation of the training sample size on the Extended Yale B database.

pixel and its neighboring ones in a local window. Although using the geodesic distance seems to be more effective than using the other distances, such as the IDLS-Distance in this experiment, ID-LARK is still weaker than IDLS, which applies the linear regression to balance the correlations between the central macro-pixel and its neighboring ones.

In the second experiment, we tested all of the algorithms on the five subsets of the Extended Yale B using the standard protocol [15], [35]. Some details of these five subsets are given in Table IV. For each person, we select one face image (i.e. the image with the most natural lighting source) from Subset 1 for training. The remaining images of Subset 1 and all images of the other subsets are for testing. The experimental results are shown in Fig. 14. From this figure, we can see that the performances of feature representation methods (with the normalization technology PP [15]) degrade with the increasingly extreme illumination for all subsets. For Subsets 1 and 2, all methods achieve very good results (100%). For the other subsets, however, our method always achieves the best results among all methods. These results demonstrate again that IDLS is more robust to lighting changes than the other methods.

TABLE IV
FIVE SUBSETS OF EXTENDED YALE B

Subsets	1	2	3	4	5
Lighting angle	$0 \sim 12$	$13 \sim 25$	$26 \sim 50$	$51 \sim 77$	> 77
Number of images	$7 \times 38 = 266$	$12 \times 38 = 456$	$12 \times 38 = 456$	$14 \times 38 = 532$	$19 \times 38 = 722$

TABLE V
THE AVERAGE RECOGNITION RATE (PERCENT) AND STD OF PCA, FLDA, LARK, ID-LARK, LBP+CHI2, LTP+CHI2, LTP/DT, GABOR+FLDA, IDLS-DISTANCE, AND IDLS WITH THE NN CLASSIFIER ON THE PIE DATABASE

Methods	5	10	15	20	25
PCA	35.9 ± 3.73	53.2 ± 4.83	61.1 ± 7.84	69.2 ± 7.05	78.4 ± 4.67
FLDA	59.3 ± 6.56	80.9 ± 4.93	86.7 ± 3.33	89.4 ± 3.26	90.2 ± 2.76
LARK	59.5 ± 6.76	74.6 ± 7.32	82.1 ± 4.63	87.7 ± 5.27	88.5 ± 3.16
ID-LARK	66.9 ± 8.24	89.1 ± 3.67	93.0 ± 2.42	95.4 ± 2.37	96.5 ± 2.00
LBP+Chi2	72.8 ± 4.85	84.0 ± 4.56	90.2 ± 3.38	93.3 ± 3.05	94.8 ± 2.63
LTP+Chi2	81.1 ± 4.64	90.3 ± 3.41	94.0 ± 2.78	96.5 ± 1.82	97.1 ± 1.37
LTP/DT	81.5 ± 4.95	90.4 ± 3.71	94.5 ± 2.43	96.6 ± 1.98	97.4 ± 1.53
TPLBP+Chi2	75.2 ± 6.64	88.8 ± 3.99	93.1 ± 3.12	96.3 ± 2.05	97.4 ± 1.51
Gabor+FLDA	75.6 ± 6.93	92.9 ± 2.32	95.4 ± 2.08	96.7 ± 2.00	97.3 ± 1.56
IDLS-Distance	68.1 ± 7.75	87.2 ± 3.24	90.8 ± 3.19	92.7 ± 2.90	93.1 ± 2.33
IDLS	82.6 ± 6.03	94.3 ± 3.89	96.3 ± 2.86	97.4 ± 2.25	97.6 ± 1.91

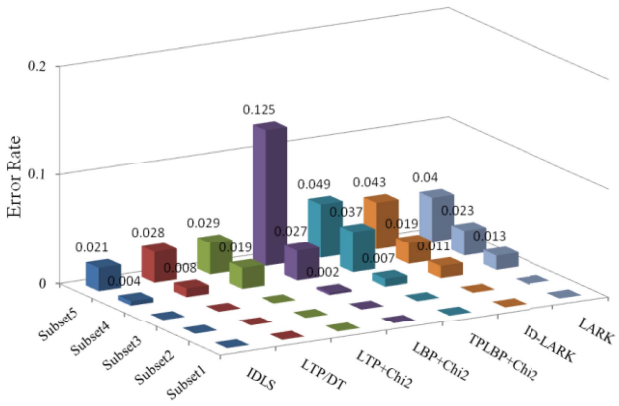


Fig. 14. The error rates on the five Extended Yale B subsets with the various feature representation methods.

Fig. 15. Sample images of a person under different illuminations and five poses in the PIE database.

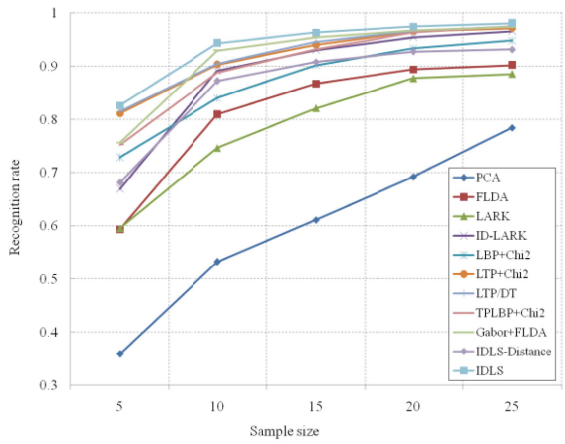


Fig. 16. The average recognition rates of PCA, FLDA, LARK, ID-LARK, LBP+Chi2, LTP+Chi2, LTP/DT, TPLBP+Chi2, Gabor+FLDA, IDLS-Distance and IDLS versus the variation of the training sample size on the PIE database.

D. Experiment Using the CMU PIE Database

The CMU PIE face database contains 68 subjects with 41368 face images as a whole [28]. Images of each person were taken across 13 different poses, under 43 different illumination conditions, and with 4 different expressions. We choose the subset of PIE database containing 50 face images (a nearly frontal pose) of each person under five different poses (there are 10 face images of each pose) [29]. All images have been cropped and resized to be 64×64 pixels. Some sample images of one person are shown in Fig. 15.

We randomly select K (K is 5, 10, 15, 20, 25) images of each class for training, and the remaining images for testing. For each K, we perform 10 runs of tests for each method with NN classifier. The average recognition rates and *std* of each method are listed in Table V. The recognition rate curve versus the variation of training sample size is shown in Fig. 16. From Fig. 16, we can see that our method IDLS achieves the best performance among all methods, irrespective of the variations of training sample size. Specifically, IDLS is significantly better than the other methods when the training sample is small. The recognition rate of IDLS is about 7%, 10% and 15% higher than that of Gabor, LBP and ID-LARK when the training sample number is 5. However, Gabor and ID-LARK can obtain the similar results with IDLS when there are enough training samples.

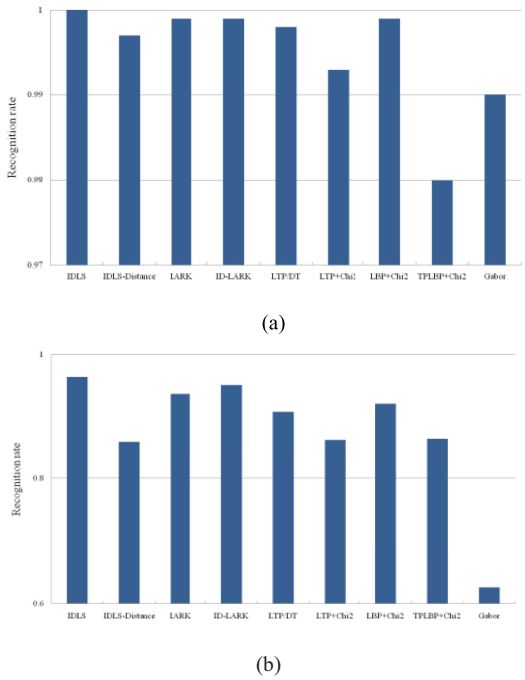


Fig. 17. The recognition rates with different feature representation methods on the PIE database under 21 different illumination conditions. (a) With the background lighting on. (b) With the background lighting off.

Additionally, we perform experiments using the protocol as adopted in [34], [35]. Two subsets are involved to test our methods. One subset contains images with the background lighting on, and other contains images with the background lighting off. Each subset includes 1428 (68×21) frontal face images under 21 different lighting conditions. For each subset, the images with the most natural lighting sources are used as gallery sets, and all of the remaining images are used as probe set. The experimental results of different feature representation methods are shown in Fig. 17. From this figure, we can see that all the methods give relatively higher recognition rates on the subset with the background lighting on, and lower recognition rates on the subset with the background lighting off. Gabor features are sensitive to illumination changes. ID-LARK achieves comparable results with LARK. The proposed method IDLS consistently achieves better results than the other methods, irrespective of whether the background lighting is on or off. These results further demonstrate that IDLS is robust to lighting changes.

E. Experiment Using the FERET Database

The FERET database [38] is widely used to test and evaluate state-of-the-art face recognition algorithms. In the FERET standard subset, the basic gallery *fa* contains 1,196 face images. the *fb* probe set contains 1,195 images of subjects taken at the same time as the gallery images but with different facial expression; the *fc* probe set contains 194 images of subjects under significantly different lighting conditions; the *Dup I* probe set contains 722 images of subjects taken between one minute and 1,031 days after the gallery image was taken; the *Dup II* probe set is a subset of the *Dup I* set, containing

Fig. 18. Face examples from the FERET database.

234 images taken at least 18 months after the gallery images. In this experiment, the face portion of each original image is automatically cropped based on the location of eyes and resized to an image of 80×80 pixels. There are 900 frontal face images to form the training sample set. Fig. 18 illustrates the cropped face examples.

We test the proposed method IDLS and compare it with the state-of-the-art methods such as LBP [17], LGBP [18], HGPP [21] and GVLBP [20]. These results of these methods have been obtained on the standard FERET probe sets. LGBP enhances the representation power due to the fact that it combines local intensity distribution with the spatial information. Actually, LGBP not only utilizes the characteristic of Gabor wavelets, but also preserves the competitive advantages of LBP. HGPP encodes phase variation using orientation change of the Gabor wavelet at a given scale, and concatenates the Gabor phase information of local neighborhoods. Additionally, GVLBP explores the information in the spatial domain among different scales and orientations.

Table VI lists the recognition rates of different methods on FERET database. From Table VI, we can see that IDLS outperforms the other methods on three probe sets (*fb*, *Dup I* and *Dup II*), although it underperforms HGPP and GVLBP on the *fc* probe set. IDLS achieves the best result in the average sense. These experimental results demonstrate the stable performance of IDLS for face recognition since the FERET database contains many variations happening in face recognition, such as expression, lighting condition and time changes.

F. Experiment Using the LFW Database

The LFW [39] database contains 13,233 target face images. There are 5,749 different individuals in the LFW. 1,680 people have two or more face images. The remainder 4,069 persons have just only one image. These images have a large degree of facial expression, occlusions, pose and illuminations since all of them are taken from the real world. Here, we use the aligned version of images[11] [40] and simply crop the face image to remove the background, leaving a 100×80 face image.

In our experiment, we follow the standard evaluation protocol as specified in [39] and compare the proposed IDLS with Gabor, LTP, LBP, TPLBP, ID-LARK and LARK for face verification in the unsupervised setting where we don't use any training samples. Note that ID-LARK and IDLS do not use the dimensionality reduction technology here. The ten subsets of the LFW database View2 are used for testing. We report the performance of each method on ten separate experiments in a

[11] http://www.openu.ac.il/home/hassner/data/lfwa/.

TABLE VI
THE RECOGNITION RATES (PERCENT) OF THE DIFFERENT METHODS ON THE FERET DATABASE

Methods	Image Size	FERET Probe Sets				
		fb	*fc*	*Dup I*	*Dup II*	*Average*
LARK	80 × 80	86.3	81.4	46.8	29.1	60.9
ID-LARK	80 × 80	91.2	93.8	45.8	22.6	63.4
Gabor+FLDA	80 × 80	93.7	83.5	51.9	25.6	63.9
LTP	80 × 80	94.9	54.6	63.8	48.8	65.5
LTP/DT	80 × 80	93.6	85.1	62.9	43.6	71.3
LBP [17]	128 × 128	97.0	79.0	66.0	64.0	76.5
LGBP [18]	88 × 80	98.0	97.0	74.0	71.0	85.0
HGPP [21]	128 × 128	97.6	**98.9**	77.7	76.1	87.6
GVLBP [20]	88 × 80	98.08	98.45	80.89	81.20	89.9
IDLS	48 × 48	**98.2**	97.0	**87.4**	**82.5**	**91.3**

TABLE VII
THE MEAN VERIFICATION RATES OF GABOR, LTP, LBP, TPLBP, ID-LARK, LARK, AND IDLS ON THE LFW DATABASE VIEW 2 (THE UNSUPERVISED SETTING)

Methods	Performance
Gabor [40]	68.47
LTP	69.95
LBP[40]	68.24
TPLBP[40]	69.26
ID-LARK	67.92
LARK[37]	72.23
IDLS	73.70

leave-one-out cross validation scheme. For LTP, ID-LARK and IDLS, we use the Euclidian distance to compute the similarity score. Table VII shows the mean verification rate of each method on the LFW database View 2. It's clear from Table VII that IDLS still achieves the best verification rate among all methods. These results further demonstrate the robustness of IDLS in real-world application.

G. Effects on IDLS With Different Parameters

In this section, we will discuss the influence of parameter setting of the scale size, macro-pixel size and down-sample factor. Here, we use different neighbors to stand for different scales as shown in Fig. 2. The performances of IDLS with different scales and macro-pixels are evaluated on the face image databases mentioned above. Experimenting on the NUST-RWFR database, the first session of each class is used for training and the rest for testing. 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 first 8 images per class of Extended Yale B database are chosen for training, and the rest for testing. In addition, the first 4 images of each pose per class are selected for training, and the remaining images for test on the CMU PIE database. Finally, we also performed experiments on the FERET database.

The recognition rates of IDLS with respect to different neighbors and sizes of macro-pixels on the five databases are shown in Fig. 19. Meanwhile, the total running time of IDLS on these databases is shown in Fig. 20. From Fig. 19, we can

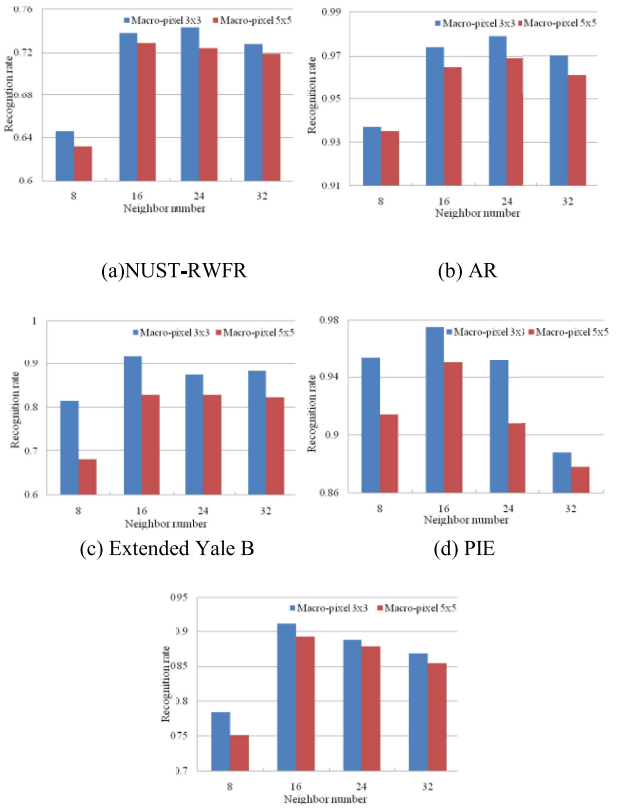


Fig. 19. The performance of IDLS with different neighbors and size of macro-pixels on standard databases.

see that a smaller size of macro-pixel (3 × 3) always achieves better results than do larger macro-pixel (5 × 5). From Fig. 20, we can see that larger macro-pixel (5 × 5) always consume more computational time than smaller sized macro-pixel (3 × 3). Therefore, we choose the size of macro-pixel as 3 × 3. With this condition, we choose $\text{IDLS}_{16,3}$ (16 means the neighbor number and 3 is the macro-pixel size) in our experiments due to the fact that it obtains a better recognition rate and costs relatively less computational time. Additionally, we also carry out the aforementioned experiments to validate the choice of the optimal down-sampling factor ($\lambda = 2$) as shown in Fig. 21.

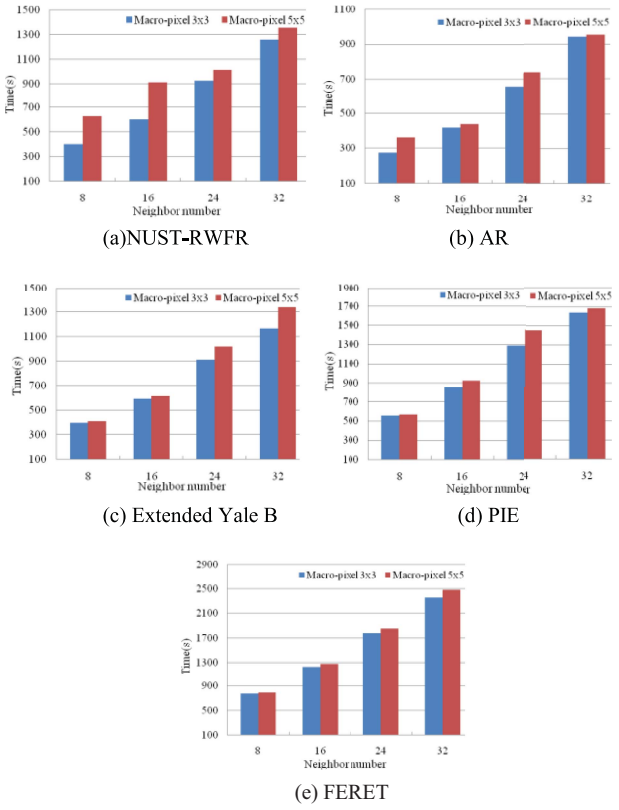


Fig. 20. The entire running time of IDLS with different neighbors and macro-pixels on standard databases.

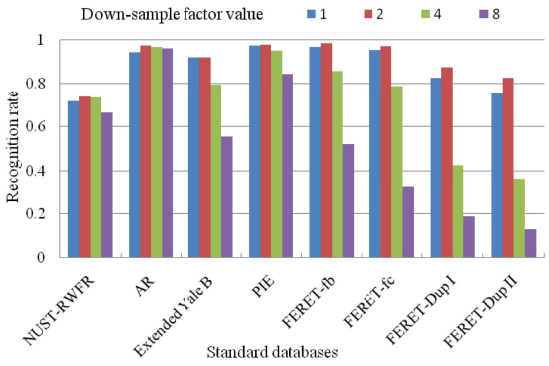


Fig. 21. The recognition rates of $IDLS_{16,3}$ using different down-sample factors.

IV. Conclusion and Future Work

In this paper, we develop a novel image feature extraction method coined IDLS and apply it to face recognition. The local structural information of IDLS is exploited by measuring the relationship between the macro-pixel of a central pixel and those of its neighbors with linear regression coefficients. One image is decomposed into a series of structure images according to local structural information. Each structure image accumulates most of the significant features in one direction. All the structure images, after being down-sampled for dimensionality reduction, are concatenated into one super-vector. FLDA is finally used to obtain the compact and low-dimensional representation for further improving the discriminative power of the method. IDLS turns out to be more robust to illumination variations than Gabor features. Our experimental results show that IDLS achieves better or comparable results in comparison with state-of-the-art methods on six face image databases.

One drawback of IDLS is that IDLS consumes more CPU time than the other methods for feature extraction, since it needs to solve a ridge regression problem once for computing the local structural information of each pixel in an image. We address this problem by resizing the image into a smaller size before implementing IDLS in this paper. Fortunately, IDLS can still yield better results using the low-resolution face images than the state-of-the-art methods using the high-resolution images. However, it is still an interesting problem as to how to extract the local structural information more efficiently. We will explore computationally more efficient algorithms for implementing IDLS, e.g. developing faster algorithms for solving the ridge regression problem, in the future.

References

[1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition: Ac literature survey," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–459, Dec. 2003.

[2] J. P. H. P. N. Belhumeur and D. J. Kriengman, "Eigenfaces versus Fisherfaces: Recognition using class specific linear projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[3] J. M. Lee, C. K. Yoo, S. W. Choi, P. A. Vanrolleghem, and I.-B. Lee, "Nonlinear process monitoring using kernel principal component analysis," *Chem. Eng. Sci.*, vol. 59, no. 1, pp. 223–234, Jan. 2004.

[4] S. Mika, G. Ratsch, J. Weston, B. Scholkopf, A. Smola, and K.-R. Muller, "Constructing descriptive and discriminative nonlinear features: Rayleigh coefficients in kernel feature spaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 5, pp. 623–628, May 2003.

[5] M. H. Yang, "Kernel Eigenfaces vs. Kernel Fisherfaces: Face recognition using Kernel methods," in *Proc. 5th IEEE Int. Conf. Autom. Face Gesture Recognit.*, May 2002, pp. 215–220.

[6] J. Yang, A. F. Frangi, J. Y. Yang, A. F. Frangi, J.-Y. Yang, D. Zhang, and Z. Jin, "KPCA plus LDA: A complete kernel fisher discriminant framework for feature extraction and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 230–244, Feb. 2005.

[7] S. C. Yan, D. Xu, and B. Y. Zhang, "Graph embedding and extensions: A general framework for dimensionality reduction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 1, pp. 40–51, Jan. 2007.

[8] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, no. 5500, pp. 2319–2323, Dec. 2000.

[9] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, Dec. 2000.

[10] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural Comput.*, vol. 15, no. 6, pp. 1373–1396, Jun. 2003.

[11] X. F. He, S. C. Yan, Y. X. Hu, P. Niyogi, and Z. Hong-Jiang, "Face recognition using Laplacianfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 3, pp. 328–340, Mar. 2005.

[12] J. Yang, D. Zhang, J. Y. Yang, and B. Niu, "Globally maximizing, locally minimizing: Unsupervised discriminant projection with applications to face and palm biometrics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 4, pp. 650–664, Apr. 2007.

[13] C. J. Liu and H. Wechsler, "Gabor feature based classification using the enhanced Fisher linear discriminant model for face recognition," *IEEE Trans. Image Process.*, vol. 11, no. 4, pp. 467–476, Apr. 2002.

[14] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.

[15] X. Y. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1635–1650, Jun. 2010.

[16] B. C. Zhang, Y. S. Gao, S. Q. Zhao, and J. Liu, "Local derivative pattern versus local binary pattern: Face recognition with high-order local pattern descriptor," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 533–544, Feb. 2010.
[17] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, Dec. 2006.
[18] W. C. Zhang, S. G. Shan, W. Gao, X. Chen, and H. Zhang, "Local gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition," in *Proc. 10th IEEE Int. Conf. Comput. Vis.*, Oct. 2005, pp. 786–791.
[19] G. Y. Zhao and M. Pietikainen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915–928, Jun. 2007.
[20] Z. Lei, S. C. Liao, M. Pietikainen, and S. Z. Li, "Face recognition by exploring information jointly in space, scale and orientation," *IEEE Trans. Image Process.*, vol. 20, no. 1, pp. 247–256, Jan. 2011.
[21] B. H. Zhang, S. G. Shan, X. L. Chen, and W. Gao, "Histogram of Gabor phase patterns (HGPP): A novel object representation approach for face recognition," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 57–68, Jan. 2007.
[22] S. Xie, S. Shan, X. Chen, and J. Chen, "Fusing local patterns of Gabor magnitude and phase for face recognition," *IEEE Trans. Image Process.*, vol. 19, no. 5, pp. 1349–1361, May 2010.
[23] M. Elad, J. L. Starck, P. Querre, and D. L. Donoho, "Simultaneous cartoon and texture image inpainting using morphological component analysis (MCA)," *Appl. Comput. Harmon. Anal.*, vol. 19, no. 3, pp. 340–358, Nov. 2005.
[24] J. L. Starck, M. Elad, and D. L. Donoho, "Image decomposition via the combination of sparse representations and a variational approach," *IEEE Trans. Image Process.*, vol. 14, no. 10, pp. 1570–1582, Oct. 2005.
[25] R. Eslami and H. Radha, "Translation-invariant contourlet transform and its application to image denoising," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3362–3374, Nov. 2006.
[26] M. J. Fadili, J. L. Starck, J. Bobin, and Y. Moudden, "Image decomposition and separation using sparse representations: An overview," *Proc. IEEE*, vol. 98, no. 6, pp. 983–994, Jun. 2010.
[27] K. C. Lee, J. Ho, and D. J. Kriegman, "Acquiring linear subspaces for face recognition under variable lighting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 5, pp. 684–698, May 2005.
[28] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 12, pp. 1615–1618, Dec. 2003.
[29] D. Cai, X. F. He, and J. W. Han, "Spectral regression for efficient regularized subspace learning," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, Oct. 2007, pp. 214–221.
[30] M. A. Turk and A. P. Pentland, "Face recognition using eigenfaces," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1991, pp. 586–591.
[31] L. Wolf, T. Hassner, and Y. Taigman, "Descriptor based methods in the wild," in *Proc. Faces Real-Life Images Workshop Eur. Conf. Comput. Vis.*, Oct. 2008, pp. 93–104.
[32] L. Wolf, T. Hassner, and Y. Taigman, "Effective unconstrained face recognition by combining multiple descriptors and learned background statistics," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 10, pp. 1978–1990, Oct. 2011.
[33] A. M. Martinez and R. Benavente, "The AR face database," CVC, Univ. Autonoma Barcelona, Barcelona, Spain, Tech. Rep. #24, Jun. 1998.
[34] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *Proc. 3rd Int. Workshop*, Oct. 2007, pp. 168–182.
[35] X. Xie, W. S. Zheng, and J. Lai, "Normalization of face illumination based on large-and small-scale features," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1635–1650, Jun. 2010.
[36] J. J. Qian and J. Yang, "A novel feature extraction method for face recognition under different lighting conditions," in *Proc. 6th Chin. Conf. Biometric Recognit.*, Dec. 2011, pp. 17–24.
[37] H. J. Seo and P. Milanfar, "Face verification using the LARK representation," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 12, pp. 1275–1286, Dec. 2011.
[38] P. J. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The FERET database and evaluation procedure for face recognition algorithms," *Image Vis. Comput.*, vol. 16, no. 5, pp. 295–306, 1998.
[39] G. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Univ. Massachusetts, Amherst, MA, USA, Tech. Rep. 07-49, 2007.
[40] L. Wolf, T. Hassner, and Y. Taigman, "Similarity scores based on background samples," in *Proc. Asian Conf. Comput. Vis.*, 2009, pp. 1–5.

Jianjun Qian received the B.S. and M.S. degrees in 2007 and 2010, respectively. Currently, he is pursuing the Ph.D. degree with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China. His current research interests include pattern recognition, computer vision, and especially face recognition.

Jian Yang (M'08) received the B.S. degree in mathematics from Xuzhou Normal University, Xuzhou, China, in 1995, the M.S. degree in applied mathematics from Changsha Railway University, Changsha, China, in 1998, and the Ph.D. degree in pattern recognition and intelligence systems from the Nanjing University of Science and Technology (NUST), Nanjing, China, in 2002. He was a Post-Doctoral Researcher with the University of Zaragoza, Zaragoza, Spain, in 2003. From 2004 to 2006, he was a Post-Doctoral Fellow with the Biometrics Centre, Hong Kong Polytechnic University, Hong Kong. From 2006 to 2007, he was a Post-Doctoral Fellow with the Department of Computer Science, New Jersey Institute of Technology, Newark, NJ, USA. He is currently a Professor with the School of Computer Science and Technology, NUST. He is the author of more than 80 scientific papers in pattern recognition and computer vision. His journal papers have been cited more than 1800 times in the ISI Web of Science, and 3000 times in the Web of Scholar Google. His current research interests include pattern recognition, computer vision, and machine learning. He is an Associate Editor of *Pattern Recognition Letters* and the IEEE TRANSACTIONS NEURAL NETWORKS AND LEARNING SYSTEMS.

Yong Xu (M'06) was born in Sichuan, China, in 1972. He received the B.S. and M.S. degrees in 1994 and 1997, respectively, and the Ph.D. degree in pattern recognition and intelligence system from the Nanjing University of Science and Technology, Nanjing, China, in 2005. He is currently with the Bio-Computing Research Center, Shenzhen Graduate School, Harbin Institute of Technology, Shenzhen, China. His current research interests include pattern recognition, biometrics, machine learning, image processing, and video analysis.