

# Local Similarity based Discriminant Analysis for Face Recognition

Xinguang Xiang<sup>1</sup>, Fan Liu<sup>2,3</sup>, Ye Bi<sup>1</sup>, Yanfang Wang<sup>2</sup> and Jinhui Tang<sup>1</sup>

<sup>1</sup>School of Computer Science and Engineering, Nanjing University of Science of Technology  
Nanjing, 210094 – CHINA

[e-mail: xgxiang@mail.njust.edu.cn, biyemail518@gmail.com, jinhuitang@mail.njust.edu.cn]

<sup>2</sup>College of Computer and Information, Hohai University, Nanjing 210098 – CHINA

[e-mail: fanliu@hhu.edu.cn, yanfang\_wang@hhu.edu.cn]

<sup>3</sup>Key Laboratory of Image and Video Understanding for Social Safety (Nanjing University of Science and Technology), Nanjing, 210094-CHINA

[e-mail: fanliu@hhu.edu.cn]

\*Corresponding author: Fan Liu

Received June 28, 2015; revised August 27, 2015; accepted September 2, 2015;  
published November 30, 2015

---

## Abstract

Fisher linear discriminant analysis (LDA) is one of the most popular projection techniques for feature extraction and has been widely applied in face recognition. However, it cannot be used when encountering the *single sample per person problem* (SSPP) because the intra-class variations cannot be evaluated. In this paper, we propose a novel method called local similarity based linear discriminant analysis (LS\_LDA) to solve this problem. Motivated by the “divide-conquer” strategy, we first divide the face into local blocks, and classify each local block, and then integrate all the classification results to make final decision. To make LDA feasible for SSPP problem, we further divide each block into overlapped patches and assume that these patches are from the same class. To improve the robustness of LS\_LDA to outliers, we further propose local similarity based median discriminant analysis (LS\_MDA), which uses class median vector to estimate the class population mean in LDA modeling. Experimental results on three popular databases show that our methods not only generalize well SSPP problem but also have strong robustness to expression, illumination, occlusion and time variation.

---

**Keywords:** Face recognition; single sample per person problem; linear discriminative analysis; local similarity

---

A preliminary conference version of this paper appeared in ACCV 2014, workshop FSLCV (Local Similarity based Linear Discriminant Analysis for Face Recognition with Single Sample per Person). This is an extended version which has been illuminated in section 1. This work was supported in part by the Natural Science Foundation of China under Grant 61301106 and 61327013, and the Natural Science Foundation of Jiangsu Province under Grant BK2012397.

## 1. Introduction

In the past two decades, plenty of research efforts have been dedicated into the domain of computer vision and pattern recognition, and tremendous progress has been achieved [1, 31-34]. Face recognition, as a nonintrusive biometric technology, has been one of the most active research topics in this field, owing to its scientific challenges and useful applications. Nowadays, appearance-based face recognition methods are the mainstream [1] and they usually represent a face image as a high-dimensional vector with plenty of redundant information. Apparently, it is necessary to extract discriminative information from such a high-dimensional vector to form a new low-dimensional vector representation. This process is called feature extraction which is beneficial to both increasing recognition accuracy and reducing computational complexity.

In the last two decades, face recognition systems are considered to be critically dependent on discriminative feature extraction, about which PCA [2,3] and LDA [4] are the two most representative ones. PCA poses the K-L transformation on training face images to find a set of optimal orthogonal bases (also called Eigenfaces) for a low-dimensional subspace and represents a face as a linear combination of Eigenfaces. Though PCA is optimal in the sense of minimum reconstruction error, it may not be favorable to classification because it does not incorporate any class specific discriminatory information into feature extraction. On the contrary, LDA takes class discriminatory information into account and seeks to find a set of optimal projection vectors by maximizing the ratio between the between-class and the within-class scatter matrices of the training samples. Existing experimental results demonstrate that LDA generally outperforms PCA in recognition rate when there are enough and representative training samples from per subject [5].

Unfortunately, in some real-world applications, the number of samples of each subject is usually very small. For example, there is only one image available in the scenario of identity card or passport verification, law enforcement, surveillance or access control. This is the so called *Single Sample per Person* (SSPP) problem which severely challenges existing feature extracting algorithms, especially their robustness to variations such as expression, illumination and disguises. For example, the performance of PCA will degrade seriously [6], while LDA even fails to work because in this case the within-class scatter matrices of the training samples will not be estimated directly.

In this paper, we propose a simple yet effective method, called local similarity based linear discriminative analysis (LS\_LDA) to address SSPP problem. The intuitive idea is that to make full use of the single image, we can divide it into many local blocks and analyze them respectively. Then based on the “divide and conquer” [18] strategy, we first classify each local block, and then integrate all the results of classification by majority voting to make a final decision. However, classifying each local block will have to face the difficulty that those local blocks corresponding to the same location but coming from different subjects might be close to each other. Intuitively, if we project these blocks from different persons into a lower-dimensional subspace and keep their respective projection in this subspace as far apart as possible, the classification will have a higher recognition rate. Although LDA is the best choice for this task, it cannot work under SSPP condition because the intra-class scatter matrices cannot be calculated. To make LDA feasible for SSPP problem, we further divide each local block into overlapped patches and propose local similarity assumption. As the patches in a local block have strong similarity, they can be regarded as to be from the same

class. Based on this idea, the within-class scatter can be computed by using the overlapped patches in a local block. Finally, the classification outputs of all the local blocks by majority voting that further improves the performance. However, considering the fact that the class sample average of existing LDA models cannot provide an accurate estimate for the class population mean when there are outliers in the training set, we further propose local similarity based median discriminant analysis (LS\_MDA) which uses the class median vector to estimate the class population mean in LDA. To evaluate the proposed methods, we perform a series of experiments on three datasets including Extended Yale B, PIE and AR face databases. Experimental results demonstrate that the proposed methods not only outperform those specially designed methods for SSPP problem, but also have strong robustness to expression, illumination, occlusion and time variation.

This paper is an extension to what we have presented earlier in our conference paper [26]. The newly incorporated work is highlighted as follows:

(1) We propose to use median discriminant analysis (MDA) to replace LDA and further improve the performance. In our previous work, we proposed LS\_LDA to solve the problem that LDA cannot work under SSPP problem. Nonetheless, the class average vector of LDA cannot provide an accurate estimate of the class population mean, especially when there are outliers. Our newly proposed LS\_MDA takes advantages of class median vector to estimate the class center and further improves the performance.

(2) We have conducted more experiments to verify the effectiveness of the proposed methods. Previously, we only provided the experimental results for SSPP problem on two databases. In this paper, the experimental results on three databases with more training samples are also published.

(3) We have further analyzed the advantages of our proposed methods and provided theoretical explanation, which were not discussed in our previous work.

The rest of this paper is organized as follows. We start by introducing related work in Section 2. Then in Section 3, we present the proposed local similarity based discriminant analysis. Section 4 demonstrates experiments and results. Finally, we conclude in Section 5 by highlighting key points of our work.

## 2. Related Work

In order to address SSPP problem, some literatures proposed to use virtual samples or generic training set. Virtual sample generation methods aim to generate extra samples for each person in order to extract the discriminatory information embedded in the intra-personal variations. For example, Shan et al. [7] presented a face-specific subspace method based on PCA which first generates a few virtual samples from single gallery image of per subject and then uses PCA to build a projection subspace for each person. In order to make LDA suitable for SSPP problem, they also extended Fisherfaces by generating virtual faces via geometric transform and photometric changes. In addition, Zhang et al. [8] applied SVD decomposition to the only face image of a person and the obtained non-significant SVD basis images were used to estimate the within-class scatter matrix of this person approximately. However, these methods are highly limited to the prior information about human face. For generic learning methods, a generic training set, in which each person has more than one training sample, is adopted to extract the discriminatory information. For example, Su et al. proposed an Adaptive Generic Learning (AGL) [24] method, which adapts a generic discriminant model to better distinguish the persons with single sample. Yang et al. proposed the sparse variation dictionary learning

(SVDL) [27] scheme by using the relationship between the gallery set and the external generic set. Recently, Deng et al. [28] proposed a novel generic learning method by mapping the intra-class facial difference of the generic faces to the zeros vectors to further enhance the generalization capability of their proposed linear regression analysis (LRA). They also proposed the extended sparse representation-based classifier (ESRC) [29] to solve SSPP problem.

All the above-mentioned methods for SSPP problem treat the whole image as a high-dimensional vector and belong to holistic representation. However, some other schemes favor local representation, in which a face image is divided into blocks and vector representation of information is conducted by blocks other than globally. Compared with holistic representation, local representation is proved to be more robust against variations [9]. For example, Chen et al. [10] proposed BlockFLD method which generates multiple training samples for each person by partitioning each face image into a set of same sized blocks and then applies FLD-based methods with these blocks. But the great differences between the appearances of long-distance blocks from one image may go against the compactness of the within-class scatter after projection. With the blocking idea, they also proposed sub-pattern based PCA (SpPCA) [22], which operates PCA directly on a set of partitioned subpatterns and extracts corresponding local sub-features and then combine them into final global features. However, decision fusion may be more effective than the feature fusion scheme of SpPCA. Recently, Zhu et al. [11] proposed patch based CRC (PCRC) for small sample size (SSS) problem. They also proposed a local generic representation (LGR) [30] based framework for SSPP problem which takes advantages of patch based local representation and generic variation representation. Kumar et al. [12] proposed patch based n-nearest classifier to improve the stability and generalization ability. Lu et al. [25] proposed a discriminative multi-manifold analysis (DMMA) method by learning discriminant feature from image patches. Generally speaking, these patch based methods has a commonality that they just consider each patch independently. Hence, they will lose the correlation information between patches, which is very important for classification.

### 3. Local Similarity based Discriminant Analysis for Face Recognition

#### 3.1 Local Similarity based Assumption

To describe local structure, we illustrate three kinds of neighborhood in Fig. 1. The  $P$  neighbor pixels on a square of radius  $R$  form a square symmetric neighbor sets. Suppose there are  $N$  pixels in an image. For the  $i$ -th pixel in the image, its  $P$  neighbor pixels can be denoted by  $\Omega_p^i = \{i_j | j = 1, \dots, P\}$ .

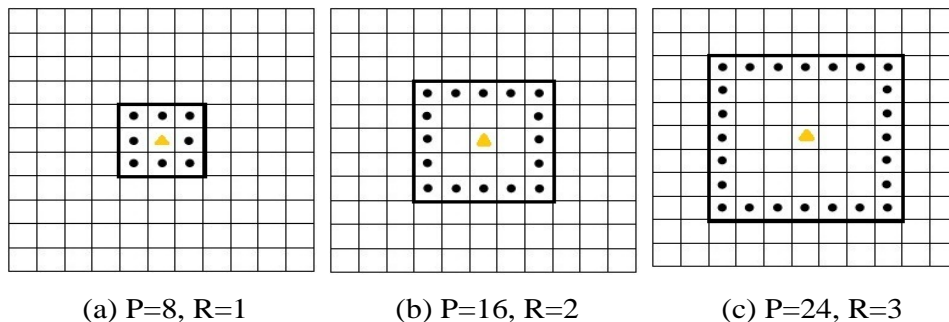
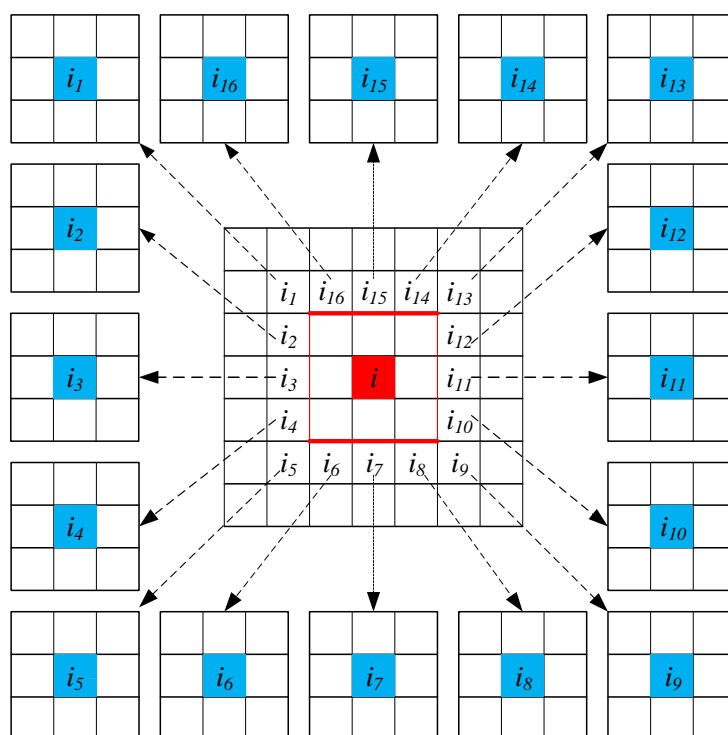


Fig. 1. Squarely symmetric neighbor sets for different R

For the  $i_j$ -th pixel in the neighbor set  $\Omega_p^i$ , it's regarded as the center of a  $S \times S$  local patch (e.g.  $S=3, 5$ ). All the  $S^2$  pixels within the patch form a  $m$  dimensional local patch vector  $\mathbf{x}_j^i, j=1, \dots, P$ , where  $m = S^2$ . Similarly, the center pixel  $i$  also corresponds to a same sized local patch which is denoted by the vector  $\mathbf{x}_0^i$ . For the  $i$ -th pixel in the image, all the local patches,  $\mathbf{x}_j^i (j=0, \dots, P)$ , determine a local block centered at the  $i$ -th pixel. **Fig. 2** shows an example of a local block which contains a central patch and 16 neighbor patches, noticing that the size of patch is  $S^2 = 3 \times 3 = 9$  and the size of the block is  $7 \times 7 = 49$ . For the pixel on the margin of an image, we use the mirror transform first and then determine its local block.



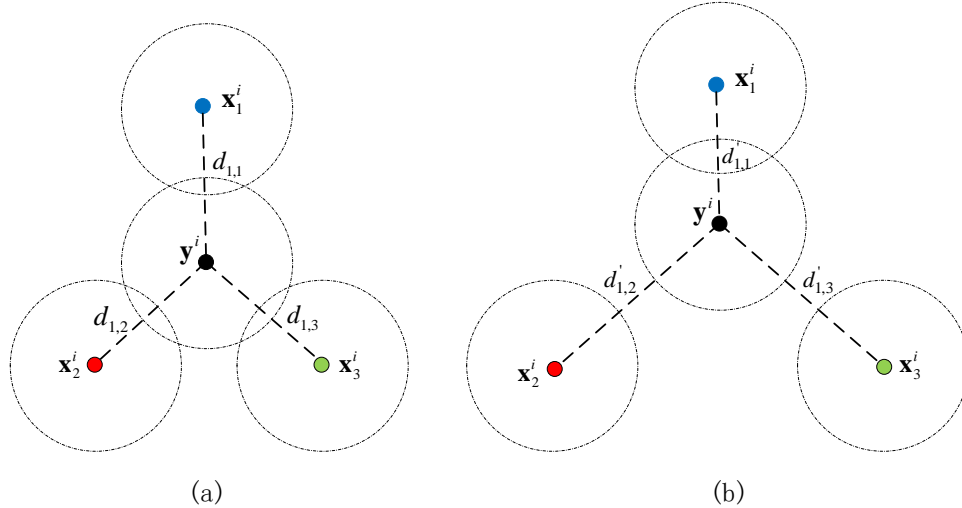
**Fig. 2.** Illustration of local patches in a local block

As the patches are overlapped and concentrate in a small block, they are strongly similar. Intuitively, they can be regarded as be from the same class. Therefore, we assume that the overlapped patches in a local block are from the same class.

### 3.2 Local Similarity based Linear Discriminant Analysis (LS\_LDA)

Suppose there are  $C$  training images  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_C\}$  and each of them belongs to a different person. According to the “dividing and conquer” [18] strategy, we first divide these training images into many local blocks. For example, the only training image  $\mathbf{x}_k$  from  $k$ -th person is divided into a set of  $N$  overlapped blocks  $\{\mathbf{x}_k^1, \mathbf{x}_k^2, \dots, \mathbf{x}_k^N\}$ , where the  $i$ -th pixel of  $\mathbf{x}_k$  corresponds to the local block  $\mathbf{x}_k^i$ . The  $i$ -th local blocks from all the  $C$  training images form a

blocking training set  $\mathbf{B}^i$ . Similarly, the probe image  $\mathbf{y}$  is also decomposed into  $N$  overlapped blocks  $\{\mathbf{y}^1, \mathbf{y}^2, \dots, \mathbf{y}^N\}$  in the same way. Then, we can classify each block respectively. However, we observe that those local blocks corresponding to the same location but coming from different subjects might be close to each other because those individuals are similar-looking.



**Fig. 3.** Illustration of the basic idea

As shown in **Fig. 3 (a)**,  $\mathbf{y}^i$  is supposed to be from Class 1 and  $\{\mathbf{x}_1^i, \mathbf{x}_2^i, \mathbf{x}_3^i\}$  are the training images' block from Class 1, Class 2, Class 3 respectively. It can be seen that the distance between  $\mathbf{y}^i$  and  $\mathbf{x}_1^i$  ( $d_{1,1}$ ) is similar with  $d_{1,2}$  and  $d_{1,3}$ . Hence, these kinds of probe image blocks are easily misclassified. To make such misclassifications less likely, we aim to learn a mapping to project these samples into a low-dimensional subspace to enlarge the between-class distance and shorten the intra-class distance, as shown in **Fig. 3 (b)**. Although LDA is the best choice for this task, it cannot work under SSPP condition because the intra-class scatter matrices cannot be calculated. To address this problem, we further divide each block into overlapped patches and take advantage of the above-mentioned local structure. As shown in **Fig. 4**, the testing block  $\mathbf{y}^i$  is divided into  $P+1$  overlapped patches, where  $\mathbf{y}^i = [\mathbf{y}_0^i, \mathbf{y}_1^i, \dots, \mathbf{y}_P^i] \in R^{m \times (P+1)}$ . Similarly, the training block  $\mathbf{x}_k^i$  can also be decomposed into  $P+1$  overlapped patches, where  $\mathbf{x}_k^i = [\mathbf{x}_{k,0}^i, \mathbf{x}_{k,1}^i, \dots, \mathbf{x}_{k,P}^i] \in R^{m \times (P+1)}$ . According to the description of local structure, these overlapped patches in local block have strong similarity. They can be theoretically regarded as to be from the same class. Then, the computation of the within-class scatter and the between-class scatter for the block training set  $\mathbf{B}^i$  is as follows:

$$\begin{aligned}
 S_w^i &= \sum_{k=1}^C \sum_{j=0}^P (\mathbf{x}_{k,j}^i - \boldsymbol{\mu}_k^i)(\mathbf{x}_{k,j}^i - \boldsymbol{\mu}_k^i)^T \\
 S_b^i &= \sum_{k=1}^C (P+1)(\boldsymbol{\mu}_k^i - \boldsymbol{\mu}^i)(\boldsymbol{\mu}_k^i - \boldsymbol{\mu}^i)^T
 \end{aligned} \tag{1}$$

Where  $\mu_k^i$  and  $\mu^i$  denote the means of the  $k$  class and all the classes respectively. Then, LDA can be used to seek the projection matrix  $\mathbf{W}$  that enlarge the between-class distance and shorten the intra-class distance in the projected subspace. It's equivalent to solve the following optimization problem:

$$J(W) = \frac{|WS_b^i W^T|}{|WS_w^i W^T|} \quad (2)$$

This optimization problem equals to the following generalized eigenvalue problem:

$$S_b^i W = \lambda S_w^i W \quad (3)$$

The solution of (3) can be obtained by applying an eigen-decomposition to the matrix  $(S_w^i)^{-1} S_b^i$  if  $S_w^i$  is nonsingular, or  $(S_b^i)^{-1} S_w^i$  if  $S_b^i$  is nonsingular. The rows of the projection matrix  $W$  are eigenvectors corresponding to the  $C - 1$  nonzero eigenvalues.

After computing the projection matrix  $W$ , we project the block training set  $\mathbf{B}^i$  and testing block  $\mathbf{y}^i = [\mathbf{y}_0^i, \mathbf{y}_1^i, \dots, \mathbf{y}_p^i] \in \mathbb{R}^{m \times (P+1)}$  into the lower-dimensional subspace and Cosine-distance based Nearest Neighbor (NN) classifier is used to classify the testing block. In order to decrease the cost of computing, we only focus on the classification of the central patch of the testing block. After classify each block, the classification outputs of all blocks can then be aggregated. Majority voting is used for the final decision making, which means that the test sample is finally classified into the class with the largest number of votes. Fig. 4 shows the whole recognition procedure, which can be summarized as “divide-conquer-aggregate” procedure.

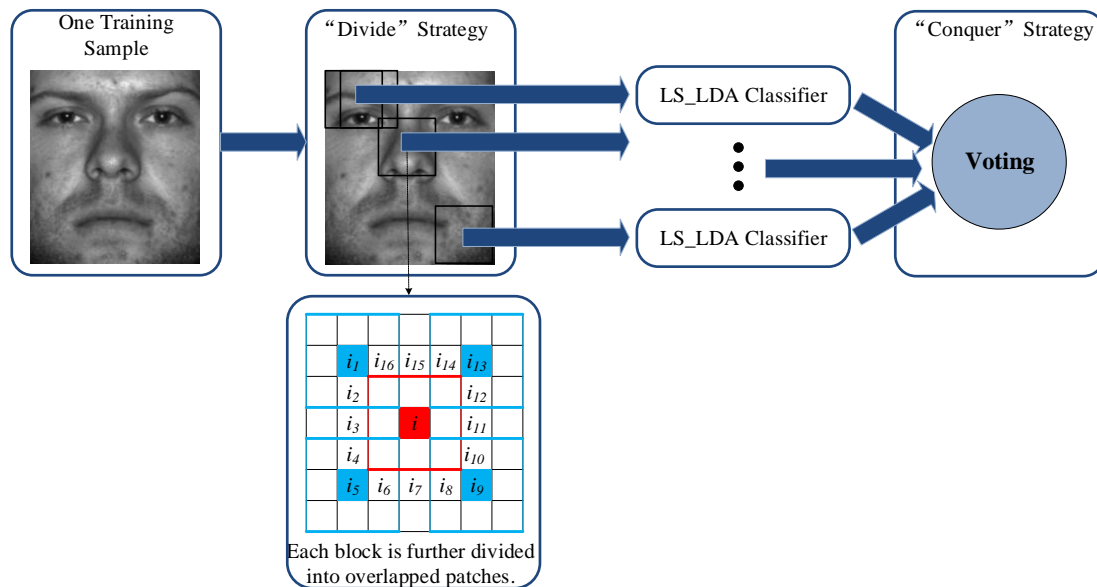


Fig. 4. The flowchart of local similarity based linear discriminant analysis for face recognition

### 3.3 Local Similarity based Median Discriminant Analysis (LS\_MDA)

In current LDA models, *class sample average* is used to estimate the *class population mean* and construct the between-class and within-class scatters. However, when the number of training samples is limited, the class sample average cannot provide an accurate estimate of the class population mean, particularly when there are outliers in the training set. To overcome this weakness, Yang et al. [20] proposed median Fisher discriminator (MFD) which uses the *class median* vector, rather than the *class sample average*, to estimate the *class population mean vector* in the LDA modeling.

In probability theory and statistics, the *median* is defined as the middle value in a distribution, above and below which lie an equal number of values. Like the sample average, the median can also be used to estimate the central tendency of the population. And, the median is more robust than the sample average, especially when there are outliers in the training sets. Specifically, given a set of  $N_k$  training sample vectors  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N_k}\}$  in Class  $k$ , we can obtain its *class median* vector  $\mathbf{M}_k$  as  $\mathbf{M}_k = (m_1, m_2, \dots, m_n)^T$ , where  $m_j$  is the median value of the  $j$ -th row in data matrix  $\mathbf{X}$ . The data matrix  $\mathbf{X}$  is defined as:

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{N_k}] = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1N_k} \\ x_{21} & x_{22} & \cdots & x_{2N_k} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nN_k} \end{bmatrix} \quad (4)$$

Then the total mean vector of all samples can be estimated by:

$$\mathbf{M} = \sum_{k=1}^C \frac{N_k}{N_{all}} \mathbf{M}_k \quad (5)$$

Where  $N_{all}$  is the number of all samples. Then the with-in class scatter and between-class scatter can be computed by:

$$\begin{aligned} \mathbf{S}_b &= \frac{1}{N_{all}} \sum_{k=1}^C N_k (\mathbf{M}_k - \mathbf{M})(\mathbf{M}_k - \mathbf{M})^T \\ \mathbf{S}_w &= \frac{1}{N_{all}} \sum_{k=1}^C \sum_{j=1}^{N_k} w_{kj} (\mathbf{x}_{kj} - \mathbf{M}_k)(\mathbf{x}_{kj} - \mathbf{M}_k)^T \end{aligned} \quad (6)$$

Where  $w_{kj}$  is the weighting coefficient, which is introduced to alleviate the influence of outliers on the construction of the within-class scatter matrix. A small weighting coefficient is imposed on the samples which are far way from the class center. Therefore the weighting coefficients  $w_{kj}$  can be generated by a monotone decreasing function with respect to the distance between the sample  $\mathbf{x}_{kj}$  and the *class median* vector  $\mathbf{M}_k$ . In [20], Yang et al. chose the Gaussian function as the weighting function:



$$w_{kj} = \exp\left(-\frac{\|\mathbf{x}_{kj} - \mathbf{M}_k\|^2}{t}\right) \quad (7)$$

Theoretically speaking, MFD is more robust to noise and outliers than LDA. The reasons lie in the following points: 1) the class median vector is more robust to estimate class center than the class sample average; 2) the weighting strategy further alleviates the influence of outliers. In small sample size cases, although there are only a few of training samples for each class, the *class median* vector  $\mathbf{M}_k$  generally provides more accurate approximation to the true class population mean, especially when there are outliers in the training set.

Inspired by MFD, we further propose local similarity based median discriminant analysis (LS\_MDA) which uses MFD to replace LDA in the LS\_LDA model. Specifically, the computation of the within-class scatter and the between-class scatter for the block training set  $\mathbf{B}^i$  is as follows:

$$\begin{aligned} \mathbf{S}_w^i &= \sum_{k=1}^C \sum_{j=1}^{P+1} w_{kj}^i (\mathbf{x}_{kj}^i - \mathbf{M}_k^i) (\mathbf{x}_{kj}^i - \mathbf{M}_k^i)^T \\ \mathbf{S}_b^i &= \sum_{k=1}^C (P+1) (\mathbf{M}_k^i - \mathbf{M}^i) (\mathbf{M}_k^i - \mathbf{M}^i)^T \end{aligned} \quad (8)$$

Where  $\mathbf{M}_k^i$  is the median vector of the  $k$ -th class from the block training set  $\mathbf{B}^i$  and  $\mathbf{M}^i$  is the total mean vector of all samples from  $\mathbf{B}^i$ .

### 3.4 Analysis of the Proposed Methods

Comparing with those global appearance based methods, the advantages of the proposed methods lie in several factors. First, the dilemma of high data dimensionality and small sample size can be alleviated by the “divide-conquer” strategy in our method, while those global appearance based methods should first use dimension reduction algorithms to solve the problem. Furthermore, each sub-image classification problem produced by “divide-conquer” strategy can be solved quickly and in parallel. Second, as local feature based methods, our proposed methods can eliminate or lower the effects of illumination changes, occlusion and expression changes by analyzing face images locally since face local information is less sensitive to those changes. However, global appearance based method are easily affected by illumination, expression changes and occlusions. Third, the local similarity assumption not only describes the relationship of the patches in a local block but also makes LDA feasible to SSPP problem. Finally, the majority voting with all the classification results of different blocks further improves the performance. It has already been found that majority voting is by far the simplest, yet as effective as more complicated schemes in improving the recognition results. This point has already been proved in [13] and [14]. In the contrary, those global appearance based methods only output a single classification results from a single classifier based on global appearance representation.

Comparing with SpPCA [22] and BlockFLD [10], our proposed LS\_LDA and LS\_MDA also utilize the blocking idea. However, we make classification for each block and combine their decisions, rather than fuse the features from each block like SpPCA [22]. Comparing with features fusion scheme, the data size of decision fusion is small. Moreover, decision fusion also has strong anti-interference capacity because the error classification results of

some blocks can be suppressed by majority voting. With regard to BlockFLD [10], although it makes LDA feasible for SSPP problem by dividing each face image into a set of same size blocks, the differences between the long-distance blocks from one image is great and enlarges the with-in class scatter after projection. LS\_LDA and LS\_MDA successfully avoid this problem by applying LDA on each block.

## 4. Experimental Results and Analysis

In this section, we use the Extended Yale B [15], PIE [19] and AR [16] databases to evaluate the proposed methods and compare them with some popular methods dealing with SSPP problem. These state-of-the-art methods include PCA [21], SpPCA [21], 2DPCA [23], BlockFLD [10], FLDA\_single [8], AGL [24], DMMA [25], and patch based CRC (PCRC) [11].

For our proposed methods, the neighbor sets are fixed at  $P = 8, R = 1$  and the patch size is fixed at  $9 \times 9$ . Since the classification of each local block can be solved independently before combing decisions, we open 12 Matlab workers for parallel computation to improve the efficiency. For the kernel parameter  $t$  in Equation 7, it's optimally chosen  $t = 4$ . Specifically, when the parameter  $t = +\infty$ , the weight  $w_{kj} = 1$  for any  $k$  and  $j$ , which means that the within-class scatter matrix is actually constructed in a non-weighted way.

### 4.1 Extended Yale B Database

The Extended Yale B face database [15] contains 38 human subjects under 9 poses and 64 illumination conditions. The 64 images of a subject in a particular pose are acquired at camera frame rate of 30 frames per 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 were used in our experiment. The cropped and normalized  $192 \times 168$  face images were captured under various laboratory-controlled lighting conditions [17] and are resized to  $80 \times 80$  in our experiments. Some sample images of one person are shown in Fig. 5.



Fig. 5. Samples of a person under different illuminations in Extended Yale B face database

To evaluate the performance of our proposed methods to SSPP problem, we use the image under the best illumination condition for training, whose azimuth and elevation are both 0 degree. The remaining 63 images are used for testing and the average results are reported. The average results of the five subsets of the Extended Yale B are also reported respectively. The details of the five subsets are given in Table 1. The experimental results are shown in Table 2. From the table, we can see that our proposed methods achieve the best average results. Moreover, LS\_MDA further improves the performance of LS\_LDA. The experimental results demonstrate that our proposed methods are robust to not only SSPP problem but also illumination variations.

**Table 1.** Five Subsets of Extended Yale B

Subsets	1	2	3	4	5
Lighting angle	0~12	13~25	26~50	51~77	>77
Number of images	6x38=228	12x38=456	12x38=456	14x38=532	19x38=722

**Table 2.** Recognition rates on Extended Yale B

Method	Subset1	Subset2	Subset3	Subset4	Subset5	Average
PCA	98.7±3.2	94.1±6.7	44.5±22.3	14.1±8.2	17.0±8.0	44.1±36.5
SpPCA	97.4±2.9	98±3.2	53.1±22.6	19.4±10.3	27±14.1	50.5±34.8
2DPCA	84.7±3.5	94.5±4.7	64.9±14.9	37.8±12.8	29.9±8.1	55.9±27.6
BlockFLD	80.3±3.2	96.3±4.1	66±7.4	59.2±16.7	53.2±11.2	67.8±19
FLDA_single	96.9±3.1	99.8±0.8	65.8±18.1	32.1±16.5	30.2±10.2	57.1±32.4
AGL	96.5±3.6	100±0	68.6±14.7	38.9±9.5	43.4±12.6	63.0±27.1
DMMA	92.5±3.1	99.1±1.7	29.8±10.7	10.3±2.4	19.9±10.8	41.7±36.6
PCRC	89±2.6	99.8±0.8	77.6±7.5	58.5±19.2	36.7±12.6	66.3±26.6
LS_LDA	<b>98.7±3.2</b>	<b>100±0</b>	94.1±6	65.8±16.2	<b>70.6±12.9</b>	<b>82.3±18.1</b>
LS_MDA	97.4±4.1	<b>100±0</b>	<b>96.5±3.0</b>	<b>86.7±12</b>	66.2±12.4	<b>85.9±16.4</b>

## 4.2 PIE Database

The CMU PIE face database contains 68 subjects with 41368 face images as a whole [19]. Images of each person were taken across 13 different poses, under 43 different illumination conditions, and with 4 different expressions. All images have been cropped and resized to be  $64 \times 64$ . Some sample images of one person are shown in Fig. 6.

**Fig. 6.** Samples of a person in PIE face database.

The images under five near frontal poses (C05, C07, C09, C27 and C29) are used in our experiment. We select 6 images with different illumination under each pose, thus we get 30 images for each individual. The single image with good illumination from frontal pose C27 is used for training. The remaining images of each subject are used for testing. The experimental result is shown in Fig. 7. We can see that LS\_LDA and LS\_MDA still achieve the best results and LS\_MDA outperforms LS\_LDA. Comparing with other methods, they leads to at least 9.3% improvements. To further evaluate the performance of our proposed LS\_LDA and LS\_MDA, we also conduct experiments with more training samples. We randomly select 1-10

samples from the 30 images of each individual for training and the remaining images are used for testing. The experiment is randomly conducted for 5 times and the average result is reported. The experimental results are shown in Fig. 8. It can be seen that LS\_MDA shows obvious advantages especially when the training sample number is small. This is because the influence of outliers is suppressed by more clean samples with the training sample number increasing.

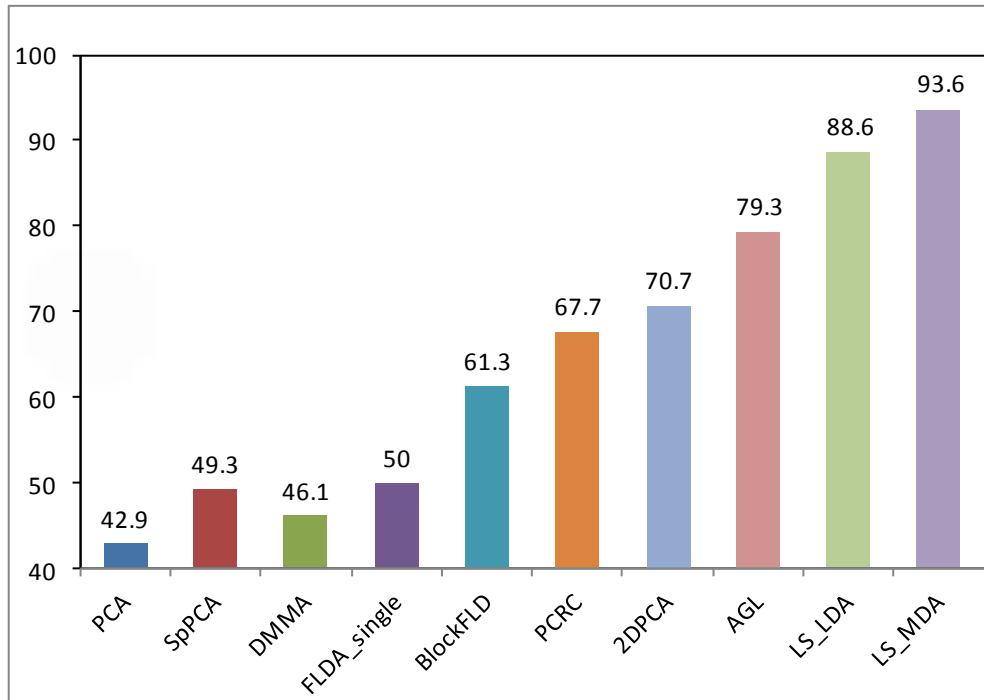


Fig. 7. Recognition rates on PIE face database.

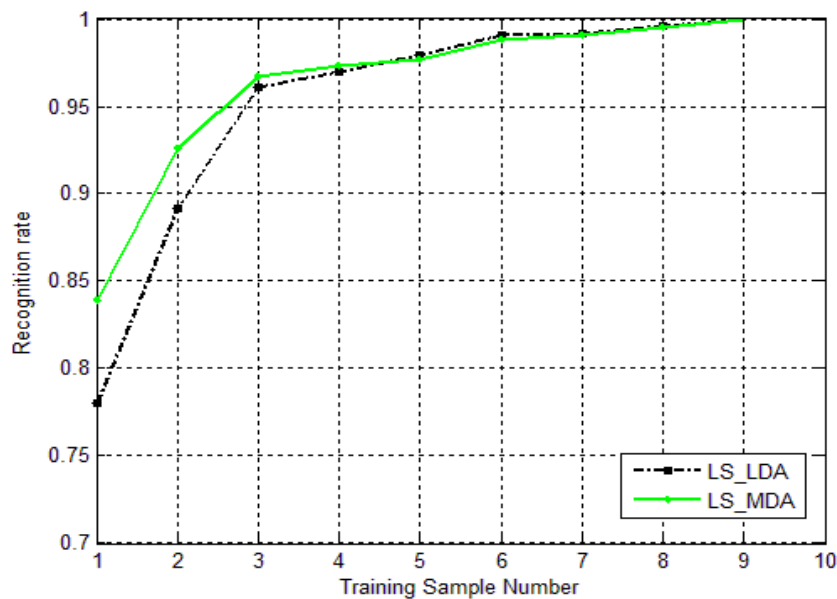


Fig. 8. Recognition rates with different training sample number.

### 4.3 AR Database

The AR face database [16] contains over 4,000 color face images of 126 people (70 men and 56 women), including frontal views of faces with different facial expressions, lighting conditions and occlusions. The pictures of 120 individuals (65 men and 55 women) were taken in two sessions (separated by two weeks) and each session contains 13 color images. To demonstrate the performance of our proposed methods for SSPP problem, we use the single image under natural expression and illumination from session 1 for training. The other images of two sessions are used for testing. These 120 individuals are selected to use in our experiment. Some samples of one person are shown in Fig. 9. The images are resized to  $32 \times 32$  and converted into gray scale.



**Fig. 9.** Samples of a person in AR database. Top row: 8 samples under different expression, illumination and occlusion from session 1; Bottom row: 8 samples from session 2 taken under the same conditions as those in top row.

The classification results on the two sessions are shown in Table 4 and Table 5 respectively. One can see from the tables that our methods show superior performance to the other methods. It's not only robust to expression and illumination variations, but also show great robustness to occlusion. In the experiments of session 1, they lead to at least 10% improvement to other methods. We also compare it with a special designed method for occlusion (GSRC [19]), which used 7 samples per subject and high resolution images but only achieved a lower recognition rate of 93% and 79% in sunglasses and scarves cases. Although there are some blocks with occlusion that will not be classified by LS\_LDA correctly, those blocks without occlusion will lead to more accurate classification results that will finally suppress the wrong results and achieve the final accurate result. In the experiments of session 2, we can find that the proposed methods are also robust to time variation. Compared to other methods, they have at least 20% improvement. To further compare the performance of the proposed LS\_LDA and LS\_MDA, we also conduct experiments with more training samples. We randomly select 1-5 samples from each subject for training and the remaining samples are used for testing. The experiment is randomly carried out for 5 times and the average results under different training sample number are shown in Table 6. From the table, we can see that LS\_MDA is superior to LS\_LDA.

**Table 4.** Recognition accuracy (%) on session 1 of AR database

Method	expression	illumination	sunglasses	scarves	average
PCA	96.4±1.7	65.3±29.9	91.1±4.2	29.7±7.3	70.6±30.6
SpPCA	95.8±1.7	86.7±10.8	90.3±1.7	40±3.6	78.2±23.8
2DPCA	97.8±2.1	93.9±2.7	95.3±1.9	56.4±4.6	85.8±18
BlockFLD	72.5±12.0	84.4±3.2	72.5±3.6	41.7±7.5	67.8±17.7
FLDA_single	93.1±3.9	96.1±13	89.4±2.9	43.1±6.5	80.4±22.9
AGL	91.4±6.7	98.9±0.5	88.3±4.6	71.7±7.1	87.6±11.4
DMMA	95.6±2.5	66.7±36.8	89.2±7.6	56.9±6.7	77.1±23.2
LS_LDA	<b>98.1±1</b>	97.2±4.8	96.9±2.1	<b>86.9±5.1</b>	94.8±5.7
LS_MDA	96.1±2.7	<b>100±0</b>	<b>97.2±1.3</b>	86.4±3.9	<b>94.9±5.8</b>

**Table 5.** Recognition accuracy (%) on session 2 of AR database

Method	expression	illumination	sunglasses	scarves	average
PCA	58.9±4.8	35±16.2	50.8±3	18.1±5.7	40.7±18.1
SpPCA	60.3±4.2	49.7±7.9	50.3±5.4	23.9±2.6	46.0±17.8
2DPCA	61.1±5.6	55.8±7.3	57.2±1.3	29.2±6.0	50.8±14
BlockFLD	33.06±12.1	41.1±4.6	30.7±5.8	22.5±2.2	31.7±9.3
FLDA_single	57.2±8.0	53.3±1.4	46.1±8.7	20.0±3.6	44.2±16.1
AGL	62.5±8.0	67.2±4.2	55±1.4	40.3±6.3	56.3±11.7
DMMA	60±7.4	39.7±23	49.7±6.5	25.8±2.2	43.8±17
PCRC	43.9±9.4	48.3±3.8	38.3±8.2	32.5±5.2	40.8±8.6
LS_LDA	<b>76.7±9.7</b>	73.6±14.8	68.1±13.8	60.3±12.1	69.7±12.7
LS_MDA	76.1±11.7	<b>81.7±6</b>	<b>69.4±11.1</b>	<b>61.9±6.5</b>	<b>72.3±11</b>

**Table 6.** Recognition accuracy (%) under different training sample number

Method	1	2	3	4	5
LS_LDA	71.3±4.8	84.8±8.3	92.1±4.1	95.3±3.6	97.7±1.0
LS_MDA	72.9±4.5	85.7±7.7	92.3±3.3	95.5±3.6	98±0.9

## 5. Conclusion

In this paper, we propose local similarity based linear discriminant analysis (LS\_LDA) to solve SSPP problem. Motivated by the “divide-conquer” strategy, we first divide the face into local blocks, and classify each local block, and then integrate all the classification results to make final decision. To make LDA feasible for the classification of each local block, we further divide each block into overlapped patches and assume that these patches are from the same class. This assumption not only reflects the local structure relationship of the overlapped patches but also makes LDA feasible for SSPP problem. Considering the fact that some blocks with illumination, expression or occlusion may be invalid for classification, we further propose LS\_MDA to improve the performance. Experimental results on two popular databases show that our methods not only generalize well SSPP problem but also have strong robustness to expression, illumination, occlusion and time variation. However, the proposed methods also rely on the basis that all the training and testing images are well aligned. Therefore, drastic pose variation will decrease their performance. This problem is planned to be solved in our future work.

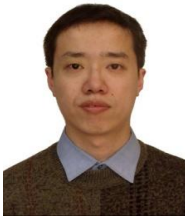
## References

- [1] W. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, “Face recognition: A literature survey,” *ACM Computing Surveys*, vol. 35, no. 4, pp. 399–458, 2003. [Article \(CrossRef Link\)](#)
- [2] M. Kirby, L. Sirovich, “Application of the Karhunen-Loeve procedure for the characterization of human faces,” *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103–108, 1990. [Article \(CrossRef Link\)](#)
- [3] M. Turk, A. Pentland, “Eigenfaces for recognition,” *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, 1991. [Article \(CrossRef Link\)](#)
- [4] P. Belhumeur, J. Hespanha, and D. Kriegman, “Eigenfaces vs. fisherfaces: Recognition using class specific linear projection,” *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997. [Article \(CrossRef Link\)](#)
- [5] A. M. Martinez and A. C. Kak, “PCA versus LDA,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 228-233, 2001. [Article \(CrossRef Link\)](#)
- [6] X. Tan, S.C. Chen, Z.H. Zhou, and h. F., “Face recognition from a single image per person: A survey,” *Pattern Recognition*, vol. 39, no. 9, pp. 1725-1745, 2006. [Article \(CrossRef Link\)](#)
- [7] S. Shan, W. Gao, and D. Zhao, “Face Identification Based on Face-Specific Subspace,” *International Journal of Imaging Systems and Technology*, vol. 13, no. 1, pp. 23-32, 2003. [Article \(CrossRef Link\)](#)
- [8] Quan-xue Gao, Lei Zhang, and David Zhang, “Face recognition using FLDA with single training image per person,” *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 726-734, 2008. [Article \(CrossRef Link\)](#)
- [9] A.M. Martinez, “Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class,” *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 748-763, 2002. [Article \(CrossRef Link\)](#)
- [10] S. C. Chen, J. Liu, Z. H. Zhou, “Making FLDA Applicable to Face Recognition with One Sample per Person,” *Pattern Recognition*, vol. 37, no. 7, pp. 1553-1555, 2004. [Article \(CrossRef Link\)](#)
- [11] P. F. Zhu, L. Zhang, Q.H. Hu, and S. C.K. Shiu, “Multi-scale Patch based Collaborative Representation for Face Recognition with Margin Distribution Optimization,” *European Conference on Computer Vision*, pp. 822–835, October 7-13, 2012. [Article \(CrossRef Link\)](#)
- [12] R. Kumar, A. Banerjee, B.C. Vemuri, H. Pfister, “Maximizing all margins: Pushing face recognition with kernel plurality,” *IEEE International Conference on Computer Vision*, pp. 2375-2382, November 6-13, 2011. [Article \(CrossRef Link\)](#)

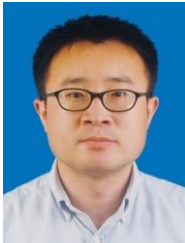
- [13] N. C. de Condorcet, *Essai sur l'Application de l'Analyze `a la Probabilit'e des D'ecisions Rendues `a la Pluralit'e des Voix*. Paris, France:Imprim'erie Royale, 1785.
- [14] L. Lam, C. Y. Suen, "Application of majority voting to pattern recognition: an analysis of its behavior and performance," *IEEE Transactions Systems Man and Cybernetics, Part A: Systems and Humans*, vol. 27, no. 5, pp. 553-568, 1997. [Article \(CrossRef Link\)](#)
- [15] A. Georghiades, P. Belhumeur, D. Kriegman, "From few to many: Illumination cone models for face recognition under variable lighting and pose," *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 6, no. 23, pp. 643-660, 2001. [Article \(CrossRef Link\)](#)
- [16] A. M. Martinez, R. Benavente, "The AR Face Database," *CVC Technical Report 24*, 1998.
- [17] K. Lee, J. Ho, and D. Kriegman, "Acquiring Linear Subspaces for Face Recognition under Variable Lighting," *IEEE Transactions Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 684-698, 2005. [Article \(CrossRef Link\)](#)
- [18] Q. F. Stout, "Supporting Divide-and-Conquer Algorithms for Image Processing," *Journal of Parallel and Distributed Computing*, vol. 4, no. 1, pp. 95-115, 1987. [Article \(CrossRef Link\)](#)
- [19] T. Sim, S. Baker, and M. Bsat, "The CMU pose, illumination, and expression database," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1615-1618, 2003. [Article \(CrossRef Link\)](#)
- [20] J. Yang, J. Yang, D. Zhang, "Median Fisher Discriminator: a robust feature extraction method with applications to biometrics," *Frontiers of Computer Science in China*, vol. 2, no. 3, pp. 295-305, 2008. [Article \(CrossRef Link\)](#)
- [21] Matthew Turk, Alex Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71-86, 1991. [Article \(CrossRef Link\)](#)
- [22] S.C. Chen, Y.L. Zhu., "Subpattern-based principle component analysis," *Pattern Recognition*, vol. 37, no. 1, pp. 1081-1083, 2004. [Article \(CrossRef Link\)](#)
- [23] J. Yang, D. Zhang, J.Y. Yang, "Two-Dimensional PCA: A New Approach to Appearance-Based Face Representation and Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131-137, 2004. [Article \(CrossRef Link\)](#)
- [24] S. Yu, S. Shan, X. Chen, W. Gao, "Adaptive Generic Learning for Face Recognition from a Single Sample per Person," *IEEE Computer Vision and Pattern Recognition*, pp. 2699-2706, June 13-18, 2010. [Article \(CrossRef Link\)](#)
- [25] J. Lu, Y.P. Tan, G. Wang, X. Zhou, "Discriminative Multi-Manifold Analysis for Face Recognition from a Single Training Sample per Person," in *Proc. of IEEE International Conference on Computer Vision*, PP. 1943-1950, November 6-13, 2011. [Article \(CrossRef Link\)](#)
- [26] F. Liu, Y. Bi, Y. Cui, Z. Tang, "Local Similarity based Linear Discriminant Analysis for Face Recognition with Single Sample per Person," in *Proc. of Asian Conference on Computer Vision, workshop FSLCV*, pp. 85-95, November 1-5, 2014. [Article \(CrossRef Link\)](#)
- [27] M. Yang, Luc Van, and L. Zhang. "Sparse variation dictionary learning for face recognition with a single training sample per person," in *Proc. of IEEE International Conference on Computer Vision*, pp. 689-696, December 3-6, 2013. [Article \(CrossRef Link\)](#)
- [28] W. Deng, J. Hu, X. Zhou, and J. Guo, "Equidistant prototypes embedding for single sample based face recognition with generic learning and incremental learning," *Pattern Recognition*, vol. 47, no. 12, pp. 3738-3749, 2014. [Article \(CrossRef Link\)](#)
- [29] Weihong Deng, Jiani Hu, and Jun Guo, "Extended SRC: Undersampled Face Recognition via Intra-class Variant Dictionary," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 9, pp. 1864-1870, 2012. [Article \(CrossRef Link\)](#)
- [30] P. Zhu, et al. "Local Generic Representation for Face Recognition with Single Sample per Person," in *Proc. of Asian Conference on Computer Vision*, pp. 34-50, November 1-5, 2014. [Article \(CrossRef Link\)](#)
- [31] Yu, R. Hong, M. Wang, and J. You, "Image clustering based on sparse patch alignment framework," *Pattern Recognition*, vol. 47, no. 11, pp. 3512-3519, 2014. [Article \(CrossRef Link\)](#)
- [32] J. Yu, D. Tao, M. Wang, and Y. Rui, "Learning to Rank Using User Clicks and Visual Features for Image Retrieval," *IEEE Transactions on Cybernetics*, vol. 45, no. 4, pp. 767-779, 2015. [Article \(CrossRef Link\)](#)



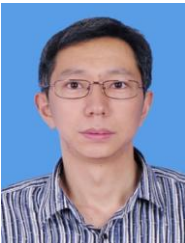
- [33] Jun Yu, Meng Wang, and Dacheng Tao, “Semisupervised Multiview Distance Metric Learning for Cartoon Synthesis,” *IEEE Transaction on Image Processing*, vol. 21, no. 11, pp. 4636–4648, 2012. [Article \(CrossRef Link\)](#)
- [34] Jun Yu, Dacheng Tao, and Meng Wang, “Adaptive Hypergraph Learning and its Application in Image Classification,” *IEEE Transaction on Image Processing*, vol. 21, no. 7, pp. 3262–3272, 2012. [Article \(CrossRef Link\)](#)



**Xinguang Xiang** received the B.A., M.S. and Ph.D. degrees in computer science from the Department of Computer Science and Technology, Harbin Institute of Technology, Harbin, China, in 2005, 2007, and 2011, respectively. He is currently with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China. His current research interests include face recognition, image analysis and understanding, video processing, and video coding.



**Fan Liu** is a lecturer in the College of Computer and Information at Hohai University, China. He received his B.S. degree and Ph.D. degree from Nanjing University of Science and Technology in 2009 and 2015, respectively. From September 2008 to December 2008, he studied at Ajou University of South Korea as an exchange student. His research interests include computer vision, image processing, pattern recognition. Dr. Liu also serves as reviewer of Information Sciences, Neurocomputing, KSII Transaction on Internet and Information Systems and Pattern Analysis and Application.



**Ye Bi** is a lecturer in the College of Computer Engineering at Huaihai Institute of Technology, China. He received his B.E. degree and M.S. degree in 1997 and 2003, respectively. From 2011, he is working for his Ph.D. degree in Pattern Recognition and Intelligence System at Nanjing University of Science and Technology. His current research interests include pattern recognition and image processing.



**Yanfang Wang** is a lecturer in the College of Computer and Information at Hohai University, China. She received her BE from Northwestern Polytechnical University, Xi'an in 2008, and her PhD from Tsinghua University, Beijing in 2014. Her current research interests are in color image enhancement, color constancy, image filtering, image super-resolution, and red-eye removal.



**Jinhui Tang** is a professor in the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China. He received his B.S. degree and Ph.D. degree from University of Science and Technology of China in 2003 and 2008, respectively. His current research interests include content-based imageretrieval, videocontentanalysis, and patternrecognition.