Local Similarity based Linear Discriminant Analysis for Face Recognition with Single Sample per Person

Fan Liu[1], Ye Bi[1], Yan Cui[2], Zhenmin Tang[1]

[1]School of Computer Science and Engineering, Nanjing University of Science and Technology, China
[2]Key Laboratory of Broadband Wireless Communication and Sensor Network Technology, Nanjing University of Posts and Telecommunications, China

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 coined 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. Experimental results on two popular databases show that our method not only generalizes well to SSPP problem but also has strong robustness to expression, illumination, occlusion and time variation.

1 Introduction

Face recognition, as a nonintrusive biometric technology, has been one of the most active research topics in the field of pattern recognition and computer vision, 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 can not be estimated directly.

In this paper, we propose a simple yet effective method, called local similarity based Linear Discriminant 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" [7] 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-dimentional 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 combination of outputs from all local blocks by majority voting further improves the performance. To evaluate the proposed method, we perform a series of experiments on two public datasets including Extended Yale B and AR face databases. Experimental results demonstrate that the proposed method not only outperforms those specially designed methods for SSPP problem, but also has strong robustness to expression, illumination, occlusion and time variation.

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 linear 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, many methods have been developed during the last two decades. Shan et al. [8] 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. But strong correlation between virtual samples decreases the representativeness of training samples and accordingly limits the performance of this method. In order to make LDA suitable for SSPP problem, Zhang et al. [9] 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. Generally speaking, the optimal number of non-significant SVD basis images is face-specific and should not be determined equally for all face images as they did. Both methods 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 [10]. For example, Chen et al. [11] 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. Recenlty, Zhu et al. [12] proposed patch based CRC (PCRC) for small sample size (SSS) problem and Kumar et al. [13] proposed patch based n-nearest classifier to improve the stability and generalization ability. However, these patch based methods has a commonality that they just consider each patch independently. Hence, they will lose the local structure information between patches, which is very important for classification. In addition, local feature can also be used to solve SSPP problem. Deng et al. [14, 15] propose two representative local feature based methods, which are uniform pursuit approach [14] and linear regression analysis (LRA)[15] respectively.

3 Local Similarity based Linear Discriminant Analysis for Face Recognition

3.1 Local Structure

To describe the local structure, we illustrate three kinds of neighborhood in Fig. 1. The neighbor pixels on a square of radius R form a squarely 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 $\Omega_P^i = \{i_j | j = 1, \ldots, P\}$.

For the i-th pixel in the image, we select a $S \times S$ local patch (e.g. S=3, 5) centered at it. All the S^2 pixels within the patch form a m dimensional local patch vector $\boldsymbol{x}_0^i$, where $m = S^2$. Similarly, the neighbor pixel i_j of the i-th also corresponds to a same sized local patch, whose patch vector is denoted by $\boldsymbol{x}_j^i, j = 1, \ldots, P$. Then , the center patch and neighbor patches determine a local block centered at the i-th pixel. Fig. 2 shows an example of a local block containing a central patch and 16 neighbor patches. The size of patch is 3×3 and the size of the block is 7×7. For the pixel on the margin of an image, we use the mirror transform first and then determine its local block.

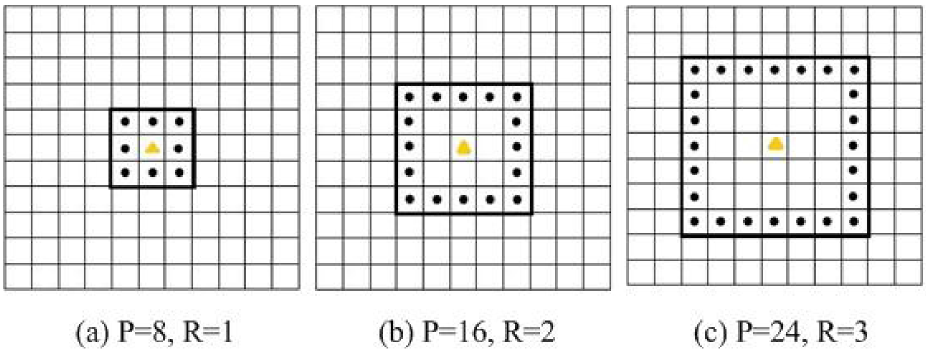

Fig. 1. Squarely symmetric neighbor sets for different R.

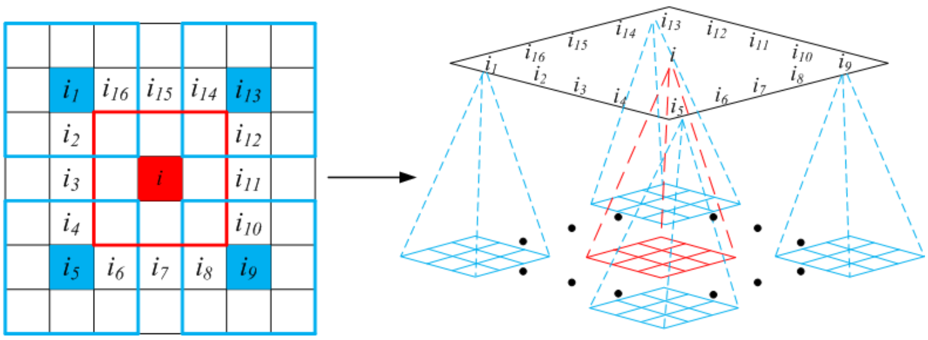

Fig. 2. Illustration of local patches in a local block.

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 $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_C\}$ and each of them belongs to a different person. According to the "dividing and conquer" strategy, we first divide these training images into many local blocks. For example, the only training image $\boldsymbol{x}_k$ from k-th person is divided into a set of N overlapped blocks $\{\boldsymbol{x}_k^1, \boldsymbol{x}_k^2, \ldots, \boldsymbol{x}_k^N\}$, where the i-th pixel of $\boldsymbol{x}_k$ corresponds to the local block $\boldsymbol{x}_k^i$. The i-th local blocks from all the C training images form a block training set $\boldsymbol{B}^i$. Similarly, the probe image $\boldsymbol{y}$ is also decomposed into N overlapped blocks $\{\boldsymbol{y}^1, \boldsymbol{y}^2, \ldots, \boldsymbol{y}^N\}$ 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.

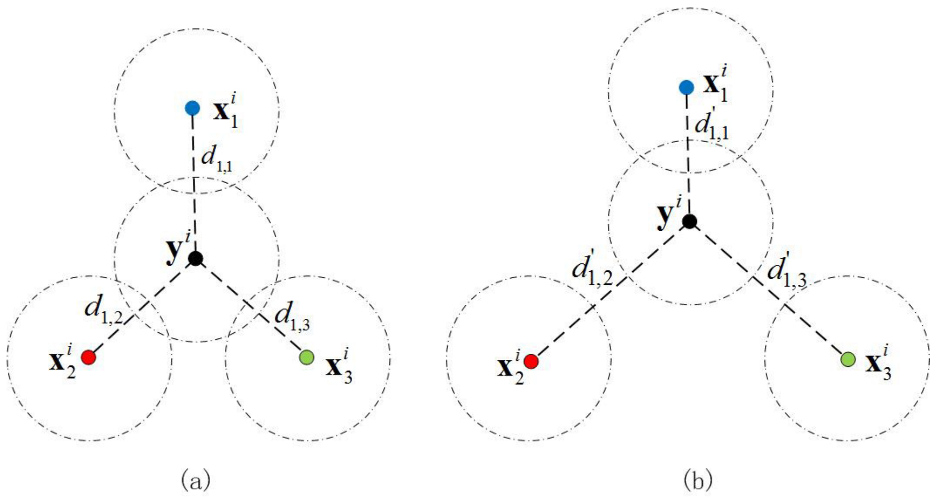

Fig. 3. Illustration of the basic idea.

As shown in Fig. 3(a), $\boldsymbol{y}^i$ is supposed to be from Class 1 and $\{\boldsymbol{x}_1^i, \boldsymbol{x}_2^i, \boldsymbol{x}_3^i\}$ are the training images' block from Class 1, Class 2 and Class 3 respectively. It can be seen that the distance $d_{1,1}$ between $\boldsymbol{y}^i$ and $\boldsymbol{x}_1^i$ 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 similarity assumption.

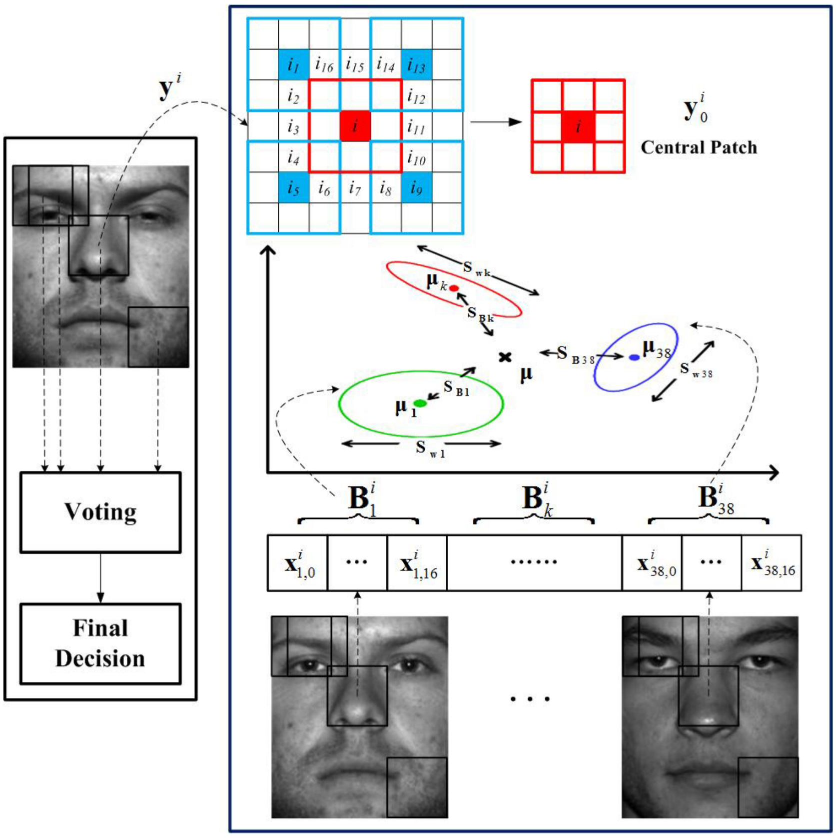

Fig. 4. Diagram of local similarity based linear discriminative and analysis for face recognition.

As shown in Fig. 4, the testing block $\boldsymbol{y}^i$ is divided into $P+1$ overlapped patches, where $\boldsymbol{y}^i = [\boldsymbol{y}^i_0, \boldsymbol{y}^i_1, \ldots, \boldsymbol{y}^i_P] \in R^{m\times(P+1)}$. Similarly, the training block $\boldsymbol{x}^i_k$ can also be decomposed into $P+1$ overlapped patches, where $\boldsymbol{x}^i_k = [\boldsymbol{x}^i_{k,0}, \boldsymbol{x}^i_{k,1}, \ldots, \boldsymbol{x}^i_{k,P}] \in R^{m\times(P+1)}$. According to the description of local similarity assumption, 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 $\boldsymbol{B}^i$ as follows:

$$S^i_w = \sum_{k=1}^{C}\sum_{j=0}^{P}(\boldsymbol{x}^i_{k,j} - \boldsymbol{\mu}^i_k)(\boldsymbol{x}^i_{k,j} - \boldsymbol{\mu}^i_k)^T \tag{1}$$

$$S^i_b = \sum_{k=1}^{C}(P+1)(\boldsymbol{\mu}^i_k - \boldsymbol{\mu}^i)(\boldsymbol{\mu}^i_k - \boldsymbol{\mu}^i)^T \tag{2}$$

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

$$J(W) = \frac{|WS_b^i W^T|}{|WS_w^i W_T|} \tag{3}$$

This optimization problem equals to the following generalized eigenvalue problem:

$$S_b^i W = \lambda S_w^i W \tag{4}$$

The solution of (4) 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.

After computing the projection matrix W, we project the block training set $\boldsymbol{B}^i$ and testing block $\boldsymbol{y}^i = [\boldsymbol{y}_0^i, \boldsymbol{y}_1^i, \ldots, \boldsymbol{y}_P^i] \in R^{m \times (P+1)}$ into the lower-dimensional subspace and 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.

4 Experimental Results

In this section, we use the Extended Yale B [16] and AR [17] databases to evaluate the proposed method and compare it with some popular methods dealing with SSPP problem. These state-of-the-art methods include PCA [3],BlockFLD [11], patch based nearest neighbor (PNN) classifier [13], FLDA_single [13] and patch based CRC (PCRC) [12].

For our proposed method, the neighbor sets are fixed at $P = 8, R = 1$ and the patch size is fixed at 11×11. Since the classification of each local block can be solved independently before combing decisions, we open 4 Matlab workers for parallel computation to improve the efficiency. For BlockFLD and PNN, the patch size is set as 10×10 and overlap is 5 pixels. In addition, histogram equalization is used as illumination preprocessing for BlockFLD, PNN and FLDA_single since it will improve their performance under illumination variation.

4.1 Extended Yale B Database

The Extended Yale B face database [16] 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 192×168 were captured under various laboratory-controlled lighting conditions [18] and are resized to 80×80 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.

Table 1. 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	$6 \times 38 = 228$	$12 \times 38 = 456$	$12 \times 38 = 456$	$14 \times 38 = 532$	$19 \times 38 = 722$

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 method achieves the best average result. It also leads to the best results on subset2, subset3 and subset4 respectively with the lighting angles increasing. However, on the subset5 which is under the worst illumination condition, our method is not superior to PNN. This is because the illumination variation is the most drastic on subset5 and more and more test blocks of each face image are regarded as "outlier" blocks that will disturb the final voting result. In addition, we also compared with classical PCA [3] since does not suffer from the SSPP problem. However, it's easy to be subject to gross variations and, thus sensitive to any changes in expression, illumination etc. Therefore, it achieves the best results on the Subset1 that is under the best illumination condition; while the recognition rate drops significantly on Subset2, Subset3, Subset4 and Subset5 whose illumination variation is drastic.

Table 2. Recognition rates (%) on Extended Yale B database for SSPP problem

Method	Subsets1	Subsets2	Subsets3	Subsets4	Subsets5	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
PNN	96.5 ± 3.6	100 ± 0	69.1 ± 15.4	43.4 ± 8.4	**75.5 ± 14.9**	73.8 ± 22.8
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
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	96.9 ± 2.6	**100 ± 0**	**95 ± 5**	**77.3 ± 14.5**	67.5 ± 12.1	**83.9 ± 16.6**

4.2 AR Database

The AR face database [17] 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. 6. The images are resized to 32×32 and converted into gray scale.

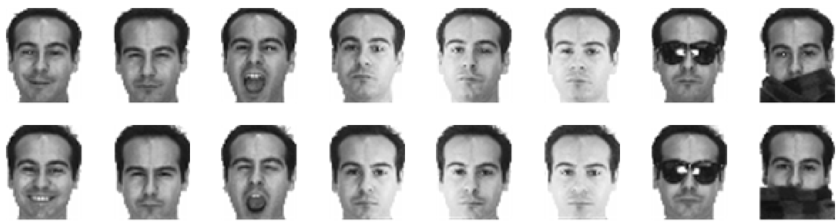

Fig. 6. Some test samples of a person in AR database. Top row: 8 samples under different expression, illumination and occlusion from session 1; Bottom row: 8 samples from session 2 taken under the same conditions as those in top row.

The classification results on the two sessions are shown in Table 3 and Table 4 respectively. One can see from the tables that our method shows superior performance to the other methods. It's not only robust to expression and illumination variations, but also shows great robustness to occlusion. In the experiments of session 1, it leads 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, there are more blocks without occlusion leading to accurate classification results. These blocks with accurate classification results will suppress the wrong results and achieve the final accurate result. In the experiments of session 2, we can find that the proposed method is also robust to time variation. Compared to other methods, it at least 20% improvement.

Table 3. Recognition rates(%) on session 1 of AR database for SSPP problem.

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
PNN	85.8 ± 10.1	91.4 ± 1.9	70.8 ± 12.9	36.9 ± 5.9	71.3 ± 23.4
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
PCRC	84.4 ± 4.9	91.4 ± 2.4	83.6 ± 3.4	63.6 ± 7.9	80.8 ± 11.7
LS_LDA	$\mathbf{96.4 \pm 2.6}$	$\mathbf{98.6 \pm 1.3}$	$\mathbf{96.1 \pm 2.1}$	$\mathbf{82.5 \pm 4.6}$	$\mathbf{93.4 \pm 7.1}$

Table 4. Recognition accuracy(%) on session 2 of AR database for SSPP problem.

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
PNN	52.5 ± 11.2	51.9 ± 4.3	34.2 ± 6.0	17.8 ± 2.1	39.1 ± 16.0
BlockFLD	33.1 ± 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
PCRC	43.9 ± 9.4	48.3 ± 3.8	38.3 ± 8.2	32.5 ± 5.2	40.8 ± 8.6
LS_LDA	$\mathbf{66.9 \pm 14.4}$	$\mathbf{74.2 \pm 9.6}$	$\mathbf{70 \pm 7.6}$	$\mathbf{56.9 \pm 12.5}$	$\mathbf{67 \pm 11.7}$

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. Experimental results on two popular databases show that our method not only generalizes well to SSPP problem but also has strong robustness to expression, illumination, occlusion and time variation. However, the proposed method also relies 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. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. Acm Computing Surveys (CSUR) **35** (2003) 399–458
2. Kirby, M., Sirovich, L.: Application of the karhunen-loeve procedure for the characterization of human faces. IEEE Trans. Pattern Anal. Machine Intell. **12** (1990) 103–108
3. Turk, M., Pentland, A.: Eigenfaces for recognition. Journal of Cognitive Neuroscience **3** (1991) 71–86
4. Belhumeur, P., Hespanha, J., Kriegman, D.: Eigenfaces vs. fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal. Machine Intell. **19** (1997) 711–720
5. Martinez, A., Kak, A.: Pca versus lda. IEEE Trans. Pattern Anal. Machine Intell. **23** (2001) 228–233
6. Tan, X., Chen, S., Zhou, Z.H., Zhang, F.: Face recognition from a single image per person: A survey. Pattern Recognition **39** (2006) 1725–1745
7. Stout, Q.F.: Supporting divide-and-conquer algorithms for image processing. Journal of Parallel and Distributed Computing **4** (1987) 95–115
8. Shan, S., Gao, W., Zhao, D.: Face recognition based on face-specific subspace. International Journal of Imaging Systems and Technology **13** (2003) 23–32
9. xue Gao, Q., Zhang, L., Zhang, D.: Face recognition using flda with single training image per person. Applied Mathematics and Computation **205** (2008) 726–734
10. Martinez, A.: Recognizing imprecisely localized, partially occluded, and expression variant faces from a single sample per class. IEEE Trans. Pattern Anal. Machine Intell. **24** (2002) 748–763
11. CHEN, S.: Making flda applicable to face recognition with one sample per person. Pattern Recognition **37** (2004) 1553–1555
12. Zhu, P., Zhang, L., Hu, Q., Shiu, S.C.K.: Multi-scale patch based collaborative representation for face recognition with margin distribution optimization. In: Computer Vision – ECCV 2012. Springer (2012) 822–835
13. Kumar, R., Banerjee, A., Vemuri, B.C., Pfister, H.: Maximizing all margins: Pushing face recognition with kernel plurality. In: 2011 International Conference on Computer Vision, IEEE (2011)
14. Deng, W., Hu, J., Guo, J., Cai, W., Feng, D.: Robust, accurate and efficient face recognition from a single training image: A uniform pursuit approach. Pattern Recognition **43** (2010) 1748–1762

15. Deng, W., Hu, J., Zhou, X., Guo, J.: Equidistant prototypes embedding for single sample based face recognition with generic learning and incremental learning. Pattern Recognition **47** (2014) 3738–3749
16. Georghiades, A., Belhumeur, P., Kriegman, D.: From few to many: illumination cone models for face recognition under variable lighting and pose. IEEE Trans. Pattern Anal. Machine Intell. **23** (2001) 643–660
17. Martinez, A.M.: The ar face database. CVC Technical Report **24** (1998)
18. Lee, K.C., Ho, J., Kriegman, D.: Acquiring linear subspaces for face recognition under variable lighting. IEEE Trans. Pattern Anal. Machine Intell. **27** (2005) 684–698
19. Yang, M., Zhang, L.: Gabor feature based sparse representation for face recognition with gabor occlusion dictionary. In: Computer Vision – ECCV 2010. Springer (2010) 448–461