

Feature Extraction Algorithm Based on Sample Set Reconstruction

Xiaoyuan Jing^{1,2}, Xianglong Ge¹, Yongfang Yao¹, Fengnan Yu¹

¹College of Automation
Nanjing University of Posts and Telecommunications
Nanjing, China
jingxy_2000@126.com

²State Key Laboratory for Software Engineering
Wuhan University
Wuhan, China

Abstract—When the number of labeled training samples is very small, the sample information people can use would be very little and the recognition rates of traditional image recognition methods are not satisfactory. However, there is often some related information contained in other databases that is helpful to feature extraction. Thus, it is considered to take full advantage of the data information in other databases by transfer learning. In this paper, the idea of transferring the samples is employed and further we propose a feature extraction approach based on sample set reconstruction. We realize the approach by reconstructing the training sample set using the difference information among the samples of other databases. Experimental results on three widely used face databases AR, FERET, CAS-PEAL are presented to demonstrate the efficacy of the proposed approach in classification performance.

Keywords—image recognition; transfer learning; sample set reconstruction

I. INTRODUCTION

Feature extraction is one of the fundamental tasks in image recognition. Also, subspace learning is an effective feature extraction and classification technique, which usually includes supervised, unsupervised and semi-supervised methods.

Linear Discriminant Analysis (LDA) [1-2] is the most representative supervised subspace learning method, the objective is to find a projection space W that maximizes the between-class scatter and minimizes the within-class scatter simultaneously. However, it is very difficult and expensive to obtain the abundant labels. For this "small samples" problem, LDA algorithm encounters several difficulties.

Transfer [3-4] is the ability to extend what one has learned to new domains. Some concepts of transfer learning [5] were discussed in a NIPS-95 workshop on "Learning to learn," which focused on the need for lifelong machine learning methods that retain and reuse previously learned knowledge. Researches on transfer learning have attached great attention since 1995. Up to now, according to the transfer levels, we can divide transfer learning into the following four aspects: (1) transfer learning based on samples [6], (2) transfer learning based on feature levels [7-8], (3) transfer learning based on parameter, (4) transfer learning based on related knowledge.

In this paper, we employ the idea of transferring the samples and present a feature extraction algorithm based on

sample set reconstruction. The recognition performance is improved by increasing the number of training samples with the assistance of transfer learning.

The rest of the paper is organized as follows: In section 2, our proposed approach, feature extraction algorithm based on sample set reconstruction is presented. Experiments are described in section 3 and section 4 concludes the whole paper.

II. FEATURE EXTRACTION ALGORITHM BASED ON SAMPLE SET RECONSTRUCTION

Generally, images in a face database may include different facial expressions or poses, under different illumination conditions, with various occlusions, and be taken in different period. In other words, the differences of face images in a database mainly reflect in the light intensity, shooting angle, pose, facial expressions and facial details etc. Inspired by the phenomenon, assuming the number of training samples is very small, we may reconstruct more samples with the difference information to help us get an optimal projection space and improve the performance of classification and recognition.

A. Sample Reconstruction

Firstly, two images of one subject are randomly selected from the AR face database and are illustrated in Figure 1:



Figure 1. two images of one subject from the AR face database

Then, several samples of one subject are chosen from the FERET database. Some difference matrices will be obtained by subtraction between each two samples in the FERET database. Then add the difference matrix to the matrices of two samples shown in Figure 1 to construct new samples shown in Figure 2.



Figure 2. new samples in the AR database reconstructed by the difference information in the FERET database

Comparing Figure 2 with Figure 1, it is notable that the difference between the original training samples in AR database and the new samples reconstructed by the information in FERET database is very small. Therefore, it is obvious to consider these new samples as training samples of the AR database to help the classification performance on the AR database.

B. feature extraction algorithm based on sample set reconstruction

Given three face datasets A, B and C, the sample number of each dataset is very small. The goal is to reconstruct training samples in one database by using the samples in other databases and achieve good performance in recognition rate. As mentioned above, we may construct the projection space of dataset A with the difference information from dataset B and C in the processing of classifying and recognizing samples on the dataset A. This section mainly focus on how to construct more correct and appropriate projections with the information of dataset B and C in order to improve the recognition performance of A.

When conducting recognition experiments on dataset A, it is considered to use the difference information of the same category in dataset B and dataset C. The procedure of the experiments on dataset B or dataset C is similar.

In order to explain the approach explicitly, the procedure is shown step-by-step:

Step 1: Select several samples belong to the same category in dataset B and dataset C respectively and get the difference data set X_B and X_C by subtracting one sample matrix from all the other matrices orderly. Samples which have the regular difference among each other are utilized in our approach. It is significant to make the most use of known category information and sample information to select samples in choosing samples from dataset B and dataset C. For example, subtracting one sample with shelter from one sample without shelter to get difference matrix has more significance.

Step 2: Reconstruct new training samples X_{Btrans} and X_{Ctrans} by using the training samples in dataset A add up or subtract the difference matrix in difference data set X_{Bdiff} and X_{Cdiff} . The number of new constructed training samples can be neither too small nor too big. On one hand, when the number is too small, it fails to construct probable projection space. On the other hand, a huge number of redundant information will influence the accuracy of the projection space.

Step 3: Train with both original training samples X_A in dataset A and new samples reconstructed in the previous after reconstructing training samples with appropriate number. Training set X , which is made up of X_A , X_{Btrans} and X_{Ctrans} , has K classes. Let X_i be the set of class i data and N_i be the number of class i data. Then a within-class scatter matrix S_w and a between-class scatter matrix S_b are defined as follows:

$$S_w = \sum_{i=1}^K S_i \tag{1}$$

$$= \sum_{i=1}^K \sum_{x_j \in X_i} (x_j - \bar{x}_i)(x_j - \bar{x}_i)^T$$

$$S_b = \sum_{i=1}^K N_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T \tag{2}$$

where S_i is the covariance of class i data; \bar{x}_i and \bar{x} are the mean vectors of class i data and all data, respectively.

The projection space W is computed by maximizing the following object function:

$$J(W) = \frac{|W^T S_b W|}{|W^T S_w W|} \tag{3}$$

Then the training samples and the test samples are projected to W to get the new training and test samples. Finally the nearest neighbor classifier is adopted for classification.

Step 4: Similar to the previous steps, B or C also can conduct the sample transferring to get more training samples.

III. EXPERIMENTS

In this section, three benchmark face databases, AR, FERET and CAS-PEAL are used to evaluate the effectiveness of our proposed method in face recognition accuracy.

Taking the current conventional computer processing power into account, we cropped every face image of these three databases to the size of 60×48 .

The AR face database [9] contains over 4000 images of 126 people (70 men and 56 women), including frontal views of faces with different facial expressions, under different lighting conditions and with various occlusions. Most of the pictures were taken in two sessions (separated by two weeks). Each session yielded 13 color images, with 119 individuals (65men and 54 women) participating in each session. We selected images from 119 individuals for use in our experiment for a total number of 3094 ($=119 \times 26$) samples. All color images are transformed into gray images with 256 gray levels. Figure.3 illustrates all of the samples of one subject, where (a)-(m) are from Session 1 and (n)-(z) are from Session 2. The details of the images are: (a) and (n), neutral expression; (b) and (o), smile; (c) and (p), anger; (d) and (q), scream; (e) and (r), left light on; (f) and (s), right light on; (g) and (t), all sides light on; (h) and (u), wearing sun glasses; (i) and (v), wearing sun glasses and left light on; (j) and (w), wearing sun glasses and right light on; (k) and (x), wearing scarf; (l) and (y) wearing scarf and left right on; (m) and (z), wearing scarf and right light on.

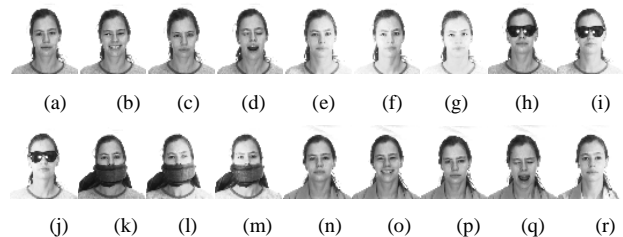




Figure 3. Demo images of one subject from the AR face database.

The FERET face database [10] is a result of the FERET program, which was sponsored by the US Department of Defense through the DARPA Program. The images were captured under various illuminations and display a variety of facial expressions and poses. Since many images in this database include the background and the body chest region, we automatically cropped every image sample. That is, we crop the image and preserve the rows from the 40th to the 340th in the original image, producing an image with size of 300×256 . We scaled the intercepted images to 60×48 . In our experiment, the database contains 2200 face images belonging to 200 persons, and there are 11 images corresponding to each person.

The details of the images are as follows: (a) regular facial status; (b) $+15^\circ$ pose angle; (c) -15° pose angle; (d) -40° pose angle; (e) -40° pose angle; (f) $+25^\circ$ pose angle; (g) -25° pose angle; (h) alternative expression; (i) different illumination; (j) $+60^\circ$ pose angle; (k) -60° pose angle. Figure 4 shows all samples of one subject.

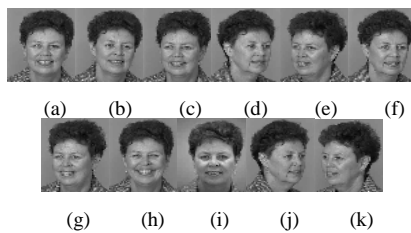


Figure 4. Demo images of one subject from the FERET face database.

The CAS-PEAL face database [11] we employed contains 1060 images of 106 individuals (10 images each person) with varying lighting. A frontal image of each subject was captured under variable illumination. Figure 5 shows a part of face images of the first person of the CAS-PEAL face database.

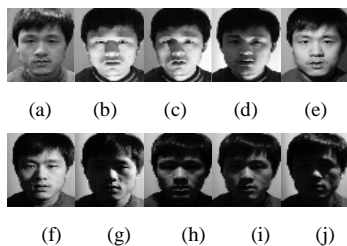


Figure 5. Demo images of one subject from the CAS-PEAL face database.

As we know, the difference between the samples of the same category in the AR database is mainly caused by expression change, illumination change and with or without shelter. The difference between the samples of the same category in the FERET database is mainly caused by angle change, expression change and illumination change. The

difference between the samples of the same category in the CAS-PEAL is mainly caused by illumination change.

Taking the experiment on the AR database as an example, the expression and illumination difference information between the samples of the same category in the FERET database and the illumination difference information between the samples of the same category in the CAS-PEAL database are used to reconstruct new samples in the AR database. Thus, the number of the training samples in the AR database is increased. Similarly, the main difference information we mentioned in the previous paragraph is utilized to reconstruct the samples in the experiments on the FERET database or the CAS-PEAL database.

TABLE I shows the recognition rates of LDA on three databases of training samples changes from 2 to 6. TABLE II shows the recognition rates of feature extraction algorithm based on sample set reconstruction on three databases of training samples changes from 2 to 6.

TABLE I. THE RECOGNITION RATES OF LDA ON AR, FERET, CAS-PEAL FACE DATABASES WHEN THE NUMBER OF TRAINING SAMPLES CHANGES FROM 2 TO 6.

Database \ Train num	Rates				
	2	3	4	5	6
AR	68.00	68.32	69.52	70.43	77.69
FERET	42.33	48.88	51.50	61.50	54.40
CAS	75.00	88.68	91.20	91.89	92.45

TABLE II. THE RECOGNITION RATES OF LDA BASED ON SAMPLE SET RECONSTRUCTION ON AR, FERET, CAS-PEAL FACE DATABASES WHEN THE NUMBER OF TRAINING SAMPLES CHANGES FROM 2 TO 6.

Database \ Train num	Rates				
	2	3	4	5	6
AR	71.16	72.62	74.39	76.43	80.71
FERET	46.25	54.72	61.03	75.64	71.29
CAS	77.85	90.14	91.96	92.87	93.62

Figure 6, Figure 7 and Figure 8 intuitively show the comparison results between our feature extraction algorithm based on sample set reconstruction and LDA algorithm on the AR, FERET and CAS-PEAL databases respectively. The red line represents for the recognition of our feature extraction algorithm based on sample set reconstruction; the green line represents for the recognition of LDA algorithm.

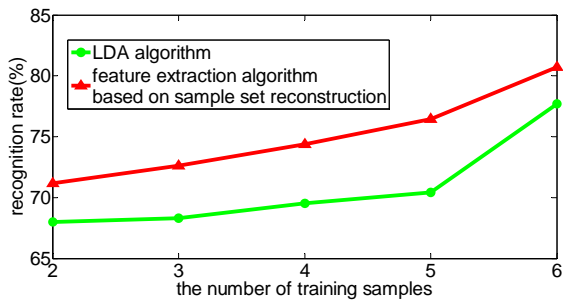


Figure 6. Comparison results between our method and LDA algorithm on AR database

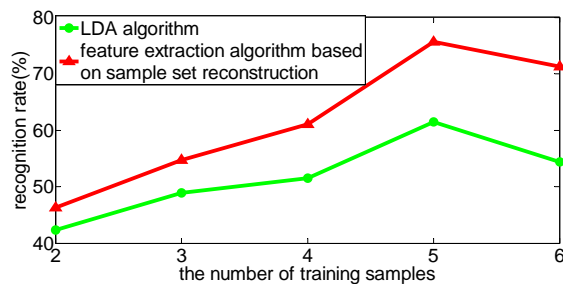


Figure 7. Comparison results between our method and LDA algorithm on FERET database

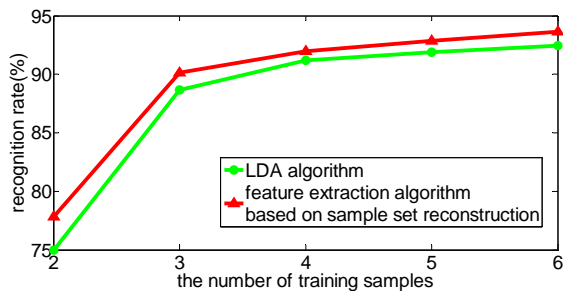


Figure 8. Comparison results between our method and LDA algorithm on CAS-PEAL database

IV. CONCLUSION

Feature extraction method based on sample set reconstruction employs the idea of transferring the samples and constructs the training sample set by taking full advantage of the difference information in other databases. The results of the experiments on AR, FERET and CAS-PEAL databases demonstrate our feature extraction

algorithm based on sample set reconstruction improves the recognition rate relative to LDA, and their performance is relatively stable.

While, there still exist some problems, such as the choice of the source domain database and the number of the reconstruction samples. We are temporarily unable to provide an optimal principle to these problems, and we hope that these problems can be solved in a follow-up work.

ACKNOWLEDGMENT

The work described in this paper was fully supported by the NSFC under Project No. 61073113 and Project No. 61272273, the New Century Excellent Talents of Education Ministry under Project No. NCET-09-0162, the Doctoral Foundation of Education Ministry under Project No. 20093223110001.

REFERENCES

- [1] H. Yu, J. Yang, "A direct LDA algorithm for high-dimensional data-with application to face recognition," *Pattern Recognition*, pp:2067-2070, 2001.
- [2] P. N. Belhumeur, J. Hespanha and D. Kiregeman, "Eigenfaces vs. Fisherfaces: Recognition using Class Specific Linear Projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, 1997.
- [3] Sinno Jialin Pan and Qiang Yang, "A survey on Transfer Learning," *IEEE Trans. Knowledge and Data Engineering*, 2010, 22(10).
- [4] D. Silver and R. Mercer, "Selective Functional Transfer: Inductive Bias From Related Tasks." In *Proc. IASTED Int. Conf. Artif. Intell. Soft Comput*, 2001, 182-189.
- [5] J. Schmidhuber, "On learning how to learn learning strategies," Technical Report FKI-198-94, Fakultat fur Informatik, 1994.
- [6] Y. Freund and R. E. Schapire, "A Decision-theoretic Generalization of On-line Learning and an Apcation to Boosting." *Journal of Computer and System Sciences*, 1997, 55(1):119-139.
- [7] Sinno Jialin Pan, James T. Kwok and Qiang Yang, "Transfer Learning via Dimensionality Reduction," In *Proceedings of AAAI*, 2008, 677-682.
- [8] Sinno Jialin Pan, Ivor W. Tsang, James T. Kwok and Qiang Yang. "Domain Adaptation via Transfer Component Analysis," *IEEE Trans. Neural Networks*, 2011, 22(2): 199-210.
- [9] Martinez, A. M., and Kak, A. C, "PCA versus LDA" *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2001, 23(2): 228-233.
- [10] P. J. Phillips, H. Wechsler, J. Huang and P. Rauss, "The FERET Database and Evaluation Procedure for Face-Recognition Algorithms," *Image and Vision Computing*. 1998, 16: 295-306.
- [11] W. Gao , B. Cao, S. Shan, X. Chen, D. Zhou, X. Zhang, D. Zhao. "The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations," *IEEE Trans. System Man, and Cybernetics (Part A)*, 2008, 38(1): 149-161.