

Personalized Face Verification System Using Owner-Specific Cluster-Dependent LDA-Subspace

Hsien-Chang Liu¹, Chan-Hung Su¹, Yueh-Hsuan Chiang¹, Yi-Ping Hung^{1,2}

¹ Department of Computer Science and Information Engineering, National Taiwan University

² Institute of Information Science, Academia Sinica, Taipei, Taiwan

E-mail: hung@csie.ntu.edu.tw

Abstract

In this paper, we propose an owner-specific cluster-dependent linear discriminant analysis (OSCD-LDA) method, and apply it to develop a personalized face verification system. Before the owner enrollment, our system first divides all the training face images into a number of clusters, each containing a subset of face images having similar characteristics. Once the owner completes the enrollment procedure, the system will assign the owner to the cluster that contains faces most similar to the owner's training faces. Then, the system uses the training faces in this most similar cluster to determine the OSCD-LDA subspace for computing the matching score. This OSCD-LDA subspace can be considered as a personalized subspace, trained specifically for this owner in order to best discriminate this particular owner from other non-owners. Our experimental results have shown that the proposed OSCD-LDA method outperforms the conventional LDA method, and can reduce false acceptance rate and false rejection rate by about 40 percent when using the XM2VTS database.

1. Introduction

How to build a reliable and convenient person authentication system is an important and active research topic. Biometric person authentication system can provide effective and efficient solutions, and has become more practical recently [4][8]. Numerous algorithms have been proposed for face recognition [2][16]. Such algorithms can be divided into two categories: geometric feature-based and appearance-based [1]. Geometric feature-based methods are robust against variations in illumination and viewpoints but very sensitive to feature extraction process.

An alternative is the appearance-based approach. Two techniques are widely used: Eigenfaces [15] and Fisherfaces [6]. Eigenfaces rely on Principal Component Analysis (PCA) and produce the most representative features, i.e., an optimal low-dimensional face representation in the least-squares sense. However, for pattern classification, it makes more sense to look for features that offer a clear

separation between the pattern classes. Hence, another method was proposed to provide the most discriminating features, which is known as Linear Discriminant Analysis (LDA) or Fisher Linear Discriminant (FLD). It has been shown that LDA outperforms PCA when large and representative training data sets are given [13].

In this paper, we propose a new method for personalized face verification, based on the *owner-specific cluster-dependent* (OSCD) LDA subspace explained below. First, we select a large number of training faces as the representative faces and partition these faces into a number of clusters. Each cluster contains a subset of face images having similar characteristics. Once the owner completes the enrollment procedure, the system will assign the owner to the cluster that contains faces *most similar* to the owner's training faces. Assume this cluster contains K clusters of faces. Next, the owner's training faces are merged into this *most similar cluster*, and the $(K+1)$ -class LDA (instead of the two-class LDA) is applied to this cluster to determine the most discriminating subspace. This subspace will be referred to as the OSCD-LDA subspace, and is a personalized subspace specifically trained for this owner in order to best discriminate this particular owner from other non-owners. It is in this personalized subspace that the matching score (or "distance") will be computed for verifying an input face. It is also in this subspace that the threshold will be determined during the on-site evaluation step.

The major contribution of this paper is that we propose to use the OSCD-LDA subspace for personalized face verification (FV). Here, the purpose of using LDA is to find a subspace in which the projected faces of different classes (including the owner and many other non-owners) are most separated from each other. The main difference between OSCD-LDA for FV and the conventional LDA for FV is that, instead of using all the training faces for LDA, we use only those training faces that are similar to the owner's faces when determining the personalized subspace. Therefore, this personalized subspace is less likely to be affected by training faces that are obviously different, and hence is more powerful in discriminating the owner from

those non-owners in the most similar cluster. (Notice that the discrimination between the owner's faces and those obviously-different faces is a relatively easy task, and can be achieved by the pre-screening described in section 2.1.3.) Our experiments have shown that the above simple design can lead to a significant improvement on the overall system performance.

2. OSCD Face Verification System

The operation of a face verification system can be divided into two stages: the training stage and the face verification stage.

2.1 Training Stage

Before the owner can use the OSCD face verification system, he needs to complete the enrollment procedure first. The training of our system consists of three steps: preliminary training, on-site training, and on-site evaluation. The last two steps are executed after the enrollment procedure is completed successfully. Figure 1 shows the flowchart of the training of our system.

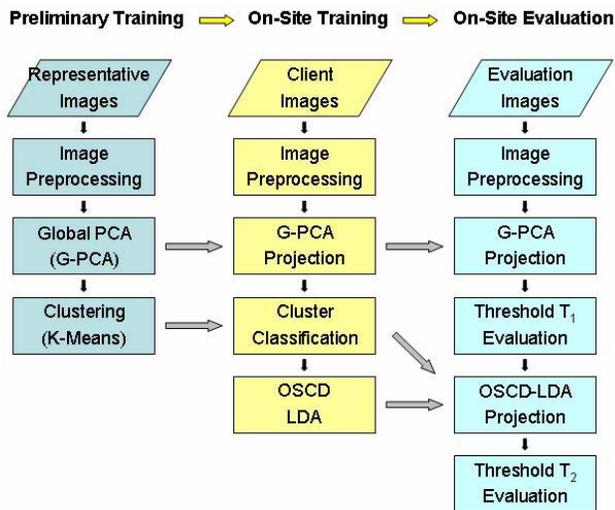


Figure 1. Flowchart of the training of the OSCD face verification system.

2.1.1 Preliminary Training

The face images selected from database are viewed as the epitome of all human face images in the world. These images are then partitioned into different clusters by using conventional clustering methods [3]. Each cluster contains a subset of face images having similar characteristics.

For this clustering, we simply adopt the widely used *K*-means algorithm. However, the clustering is not performed directly with the original face images. Instead, we first reduce the dimension of the feature space by applying a *global PCA* projection to the training images before clustering them. The reason for this dimension reduction is because the distances among original face images are usually very large in high-dimensional space and hence it is hard to obtain a good clustering result in this situation. The distance measure will be more meaningful after dimension reduction. It can be seen from Figure 2 that face images partitioned into the same cluster are more similar to each other than they are to face images in other clusters.

2.1.2. On-Site Training

After completing the enrollment procedure, the user becomes the owner. Once he becomes the owner, our system will capture a set of his face images and perform owner-specific training, which can be divided into two steps: on-site training and on-site evaluation. Some of these images will be used for on-site training, and the remaining saved for on-site evaluation.



Figure 2. Some face examples in each cluster when the number of clusters is set to be 5 (using XM2VTS database).

As shown in Figure 1, after image preprocessing, we first project the mean face of the owner onto the *global-PCA* subspace obtained from the preliminary training, and then find the *most similar cluster* that is closest to the owner's mean in the global-PCA subspace. Here, we simply use Euclidean distance between the owner's mean and the cluster's mean in global-PCA subspace:

$$S_{Ei} = \sqrt{(\tilde{\mathbf{m}}_o - \tilde{\mathbf{m}}_i)^T (\tilde{\mathbf{m}}_o - \tilde{\mathbf{m}}_i)}$$

where \mathbf{m}_o is the owner's mean face and *tilde* indicates the projection onto a subspace.

Next, the owner's training faces are merged into the *most similar cluster*, and the LDA method is applied to, and only to, this cluster to determine the OSCD-LDA subspace. This OSCD-LDA subspace will be used for computing the matching score (or "distance") when verifying an input face.

2.1.3 On-Site Evaluation

The objective of on-site evaluation is to decide the two OSCD-threshold values, T_1 and T_2 , needed in the face verification stage. Threshold T_1 is used to decide whether the input face is close enough to the *most similar cluster* in the global PCA space, which can be thought of as a pre-screening. If the input face passes the pre-screening, threshold T_2 is then used to decide whether the user is indeed the owner. Here, we need a set of images, called the evaluation image set. This image set consists of two subsets: one is owner-evaluation subset, X_{oe} , and the other is impostor-evaluation subset, X_{ie} . The impostor-evaluation subset contains face images that are different from the representative faces used in the preliminary training.

For threshold T_1 , we need only the information of the owner-evaluation set X_{oe} . After projecting X_{oe} to the global PCA-subspace, we can obtain the distances between each *projected owner-evaluation face* and the *projected mean-face of the most similar cluster*, and find the maximum distance. In our current implementation, threshold T_1 is simply set to be 1.2 times the maximum distance.

The determination of threshold T_2 is a little more complicated. The evaluation image set is first projected onto the OSCD-PCA subspace (to overcome the small sample size problem) and then subsequently projected onto the OSCD-LDA subspace. This is the space where the matching score (or "distance") will be computed later in section 2.2. Also, it is in this space that we determine T_2 to be the value that leads to FAR = FRR in our experiments.

The choice of a good matching score in the OSCD-LDA subspace is also an important issue. The Euclidean distance is the simplest choice, but its performance is usually poor. Kitter et al. have shown that the normalized correlation and the gradient direction metric can achieve much better performance [10]. We have tried all the three measures in our experiments, and choose to use the gradient direction metric for its best performance (see section 3).

2.2. Face Verification Stage

Once the above training stage is completed, our system can be put to work as a *personalized* face verification system.

When a user claims to be the owner and wants to access the machine, his face image must pass two tests. For the first test, his face image (after some pre-processing) is projected onto the global PCA subspace. The distance between the projected face and the projected mean face of the *most similar cluster* must be smaller than threshold T_1 .

If the user passes the first test, his face image will be first projected onto the OSCD-PCA subspace, and finally onto the OSCD-LDA subspace. The second test compares threshold T_2 with the "distance" between this OSCD-LDA-projected input face and the OSCD-LDA-projected mean of the owner's training faces. If the "distance" is smaller than T_2 , the identity of the user is verified as the owner; otherwise he will be asked to pass other stricter and more tedious tests.

3. Experiments

To compare the performance of our OSCD-LDA method with that of the conventional LDA method, we use the well-known XM2VTS database [14]. We adopt three different score measures in the comparison, i.e., Euclidean distance, S_E , normalized correlation, S_N , and gradient direction metric, S_D . The experiments were carried out according to the Lausanne protocol [11]. This protocol divides the database into 200 clients and 95 impostors. Each subject has eight images in total. Among the eight images, three are used as training images, three as evaluation images, and the remaining two as test images. Each of the 200 clients is treated as the owner in turns, and the remaining 199 clients are used as the epitome of all faces for training. In the test, each client will only be used to access his own personalized system, while the impostors will be used to access all the 200 personalized systems.

Tables 1 and 2 show the results obtained by using the registered XM2VTS database. The results obtained with the evaluation set and with the test set are shown separately. The evaluation set, which is independent of the training set, is used to compute the receiver operating characteristics (ROC curve) by varying the OSCD-thresholds and measuring the corresponding false rejection rate (FRR) and the false acceptance rate (FAR). The ROC curves obtained with the conventional LDA method and with our OSCD-LDA method are shown in Figures 3, respectively. The point on the ROC curve that satisfies FAR=FRR is selected as the operating threshold for the subsequent test, which test uses another independent set of test images. The results are recorded in the last two columns of Tables 1 and

2. Note that in general there is a close agreement between the results obtained with the evaluation set and the test set.

From Tables 1 and 2, it can be seen that the proposed OSCD-LDA method obviously achieves better performance than the conventional LDA method. The FAR and FRR are reduced from 4.90% to 2.91% and from 8.25% to 4.75%, which indicates a 40 percent reduction in error rates.

The operating thresholds we used in the above experiments was determined by setting FAR=FRR. However, the obtained FRR is obviously worse than the FAR. It is because XM2VTS has only eight face images for each client (and hence for the owner) and only three of them are used for evaluation. In practice, the system can have as many face images of the owner as it needs, and this will not be a problem anymore.

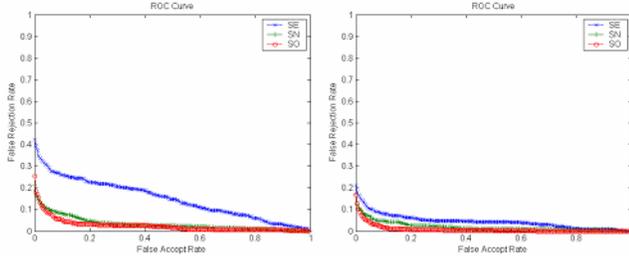


Figure 3. ROC curves obtained by using ordinary LDA (Left) and OSCD-LDA (Right).

60x72	Evaluation Set		Testing Set	
	FAR	FRR	FAR	FRR
S_E	20.66	20.33	22.49	13.25
S_N	5.85	5.00	6.62	7.75
S_O	4.53	3.00	4.90	8.25

Table 1. Experimental results obtained by using the conventional LDA method.

60x72 cluster	Evaluation Set		Testing Set	
	FAR	FRR	FAR	FRR
S_E	6.68	7.67	7.94	7.00
S_N	4.30	3.83	4.45	5.50
S_O	1.96	2.50	2.91	4.75

Table 2. Experimental results obtained by using the proposed OSCD-LDA method.

4. Conclusion

This paper proposed a new method for developing a personalized face verification system. The proposed method first partitions all the possible face images into a number of clusters, each representing a subset of face images having similar characteristics. Then, depending on which cluster the owner is associated with, the system determines the

corresponding OSCD-LDA subspace for computing the matching score for face verification. Our experiments have shown that, when using the XM2VTS database, the proposed OSCD-LDA method outperforms the conventional LDA method, and can reduce FAR and FRR by about 40 percent.

References

- [1] R. Brunelli and T. Poggio, "Face Recognition: Features versus Templates," IEEE Transactions on PAMI, vol. 15, no.10, pp. 1042-1052, 1993.
- [2] R. Chellappa, C. L. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," Proceedings of IEEE, vol. 33, no. 5, pp. 705-740, 1995.
- [3] D. S. R. Duda, and P. Hart, Pattern Classification, Wiley, New York, 2001.
- [4] J. L. Dugelay, J. C. Junqua, C. Kotropoulos, R. Kuhn, F. Perronnin, and I. Pitas, "Recent Advances in Biometric Person Authentication," Proceedings of IEEE Int Conf on ASSP, vol. 4, pp. 4060-4063, 2002.
- [6] P.N. Belhumeur, J.P. Jespanha, and D.J. Kriegman. "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," IEEE Transactions on PAMI, 19(7):711-720, July 1997.
- [7] R. Huang, Q. Liu, H. Lu, and S. Ma, "Solving the Small Sample Size Problem of LDA," Proceedings of ICPR, vol. 3, pp. 29-32, 2002.
- [8] A. K. Jain, L. Hong and S. Pankanti, "Biometric: Promising Frontiers for Emerging Identification Market," Communications of ACM, vol. 2, pp. 91-98, 2002.
- [10] Y. Li, J. Kittler, and J. Matas, "On Matching Scores of LDA-based Face Verification," Proceedings of British Machine Vision Conference, vol. 1, pp. 42-51, 2000.
- [11] J. Luetttin and G. Maître, "Evaluation Protocol for the Extended M2VTS Database (XM2VTSDB)," Technical Report IDIAP-COM 98-05, IDIAP, 1998.
- [13] A. M. Martinez and A. C. Kak, "PCA versus LDA," IEEE Transactions on PAMI, vol. 23, no. 2, pp. 228-233, 2001.
- [14] K. Messer, J. Matas, J. Kittler, J. Luetttin, and G. Maitre, "XM2VTSDB: The Extended M2VTS Database," International Conference on Audio and Video-based Biometric Person Authentication, pp. 72-77, 1999.
- [15] M. Turk and A. Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.
- [16] W. Zhao, R. Chellappa, A. Rosenfeld, and P. J. Phillips, "Face Recognition: A Literature Survey," Technical Report CAR-TR-948, CS-TR-4167, N00014-95-1-0521, University of Maryland, 2000.