LDA WITH SUBGROUP PCA METHOD FOR FACIAL IMAGE RETRIEVAL

Wonjun Hwang, Tae-Kyun Kim, Seokcheol Kee

Human Computer Interaction Lab., Samsung Advanced Institute of Technology, Korea.
wjhwang@sait.samsung.co.kr

ABSTRACT

We need to capture properties of face image variations from training persons and generalize them to a new test person for robust image retrieval, which is essential for the case that there is only a single image of the test person to retrieval face images of the equal person from database. Conventional methods of Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) can be exploited to learn robust subspaces for face image variations contained in the training face images and the subspaces can be coped with the similar variations of new test persons. In this paper, a new method of LDA with subgroup PCA is proposed for an efficient learning of face image changes. Particularly the proposed method is designed to learn many kinds of extrinsic variations that are included in the training set. For this purpose the training set is partitioned into several subgroups by local property and PCA applied to the each subgroup yields the specialized subspace representation for the variations of the subgroups. On the contrary LDA is performed on the subgroup PCA features of the whole training set in order to avoid an overfitting under the influence of dominant variations. Several representations of the LDA with the subgroup PCA are finally merged by the weighted sum rule and it is shown to yield better recognition rate in the experimental results of the fluent face images collected from Purdue, PIE, ALTKOM, XM2VTS and BANCA databases.

1. INTRODUCTION

Face recognition has been widely studied for improving biometric security and retrieving the multimedia data which contains a human face. Of many approaches[1] which have been recently proposed, LDA[2] is one of the powerful methods for face recognition and its representation maximizes the ratio of between class and within class scatters for classification. Because LDA directly applied to face images has the singularity problem at the within-class scatter matrix, PCA is commonly pre-performed for dimensionality reduction of the face image[3]. The PCA yields a low dimensional subspace which minimizes the mean square reconstruction error by finding the directions of major variations in the whole training set considered.

Face recognition performance by using statistical learning techniques of PCA and/or LDA depends on available training face images and particularly the number of images. The main key to improve the performance is in an efficient learning of properties of the limited training face images. It is noted that simply collecting more variations in the training set for PCA and LDA learning does not always give better performance. For example, the variations in illumination and pose of face images have quite different properties and except one dominant variation, others can not affect learning scheme. In such case the resampling of the training data set enhanced the performance of PCA-LDA[4]. In more detail, PCA and LDA learned from any subset can yield more efficient representation for any specific variation. Fusion of these specialized representations to each subset can enhance recognition rate. However, the system can be overfitted for local variation of the subset even though there is an integration scheme.

In this paper, we assume that various images of generic (training) faces are collected in order to learn the property of variations. Collected face images are divided into several subgroups specialized to specific kind of variation like illumination and pose changes. These subgroups are utilized for producing the several eigenface sets by PCA. LDA is then performed on the each PCA representation of the whole training data set, not subgroup. We call this method as “LDA with subgroup PCA.” In the proposed method, subgroup PCAs can discard fewer dimensions and embrace the specific kind of variation effectively. Moreover LDA of the whole training set after projecting on the subgroup eigenfaces learns more discriminative information. Finally the weighted sum rule is proposed for combining the results of the specialized classifiers of LDA with subgroup PCA. The experimental result shows the benefit of the proposed method in the varying face database.

The rest of this paper is organized as follows: Section 2 derives the LDA with subgroup PCA. In Section 3, the merging scheme of the several classifiers is explained. The experimental results and discussion are drawn in Section 4 and the conclusion is summarized in Section 5.

2. LDA WITH SUBGROUP PCA METHOD

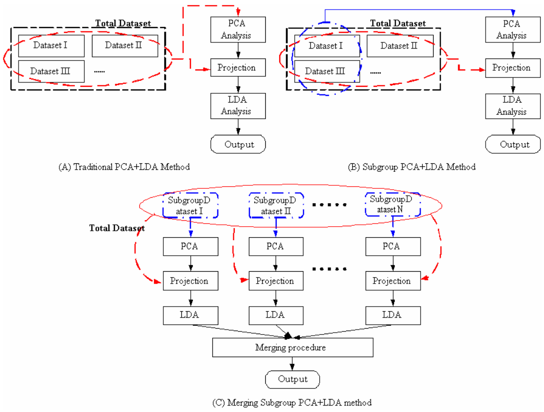


Fig. 1. Procedures of traditional PCA and LDA scheme and proposed LDA with subgroup PCA scheme

The training dataset χ is partitioned into N subgroups, $\chi = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N\}$. There are many methods for such division, for example simple K-Nearest Neighboring algorithm, but here we heuristically divided the database according to extrinsic features and each flocked subgroup has the similar property. The two subgroups considered in this paper are an illumination variation based set and a pose variation based set. The illumination and pose variations are key hurdles to disturb successful face image retrieval. After the clustering, we had N eigenface sets from each subgroup called simply "subgroup eigenface." Each set of subgroup eigenface can efficiently embrace its own local property compared with the set of eigenface learned from the entire training dataset, called as global eigenface. i th subgroup eigenface, $\mathbf{U}_i$ is derived by

$$\Sigma_i \mathbf{U}_i = \mathbf{U}_i \mathbf{\Lambda}_i, \quad (1)$$

where $i = 1, 2, \cdots, N$ is a subgroup number, Σ_i is the covariance matrix of i th subgroup, and $\mathbf{\Lambda}_i$ is the its eigenvalue matrix. The corresponding projection is obtained by

$$\alpha_i = \mathbf{U}_i^T (x - \mu_i), \quad (2)$$

where α_i is the projected vector, μ_i is the mean vector, and x is a face image vector. Hereby we obtained the N optimized eigenfaces, $\mathbf{U}_i$ for the property of each subgroup and they are exploited to reduce the dimensionality of input images.

The projected vectors are acquired from all the training face images for the subsequent LDA. In the method of PCA and LDA using resampling training dataset[4], same subsets are utilized at both stages of PCA and LDA as shown at Fig 1. (A). On the contrary, we utilized the subsets for PCA and the whole training set for LDA to avoid an overfitting and achieve more discriminant results as shown at Fig. 1 (B) and (C). The transformation matrix, W_i is derived by

$$W_i = \arg\max_W \frac{W_i^T S_B^{'} W_i}{W_i^T S_W^{'} W_i} = \arg\max_W \frac{W_i^T (U_i^T S_B U_i) W_i}{W_i^T (U_i^T S_W U_i) W_i}, \quad (3)$$

where S_B is the between-class scatter matrix, S_W is the within-class scatter matrix. The projection function of the i th LDA with subgroup PCA is derived by

$$y_i(x) = (U_i W_i)^T (x - \mu_i) = B_i^T (x - \mu_i), \quad (4)$$

where y_i is the final feature vector and B_i is the final transformation matrix. By this way, we can extract the projected vectors which are donated to variation of a subgroup and more discriminable for a new test face.

Comparing with the proposed method, the main problem of the traditional PCA and LDA is in that lots of sources of information are eliminated in the step of PCA to solve the singularity problem of LDA. Consequently LDA can not produce the best separable result with such reduced dimensionality. For example, if the training dataset consists of two groups which have different properties and the variation of one group is larger than the other, global eigenfaces are biased to the property of the large group. On the other side, LDA with subgroup PCA method can represent the various properties of training dataset regardless of sizes and variations of different groups. Basis images as shown in Fig. 2 show different properties of the subgroups definitely. For example, Fig. 2 (C) presents higher frequency features such as edges of

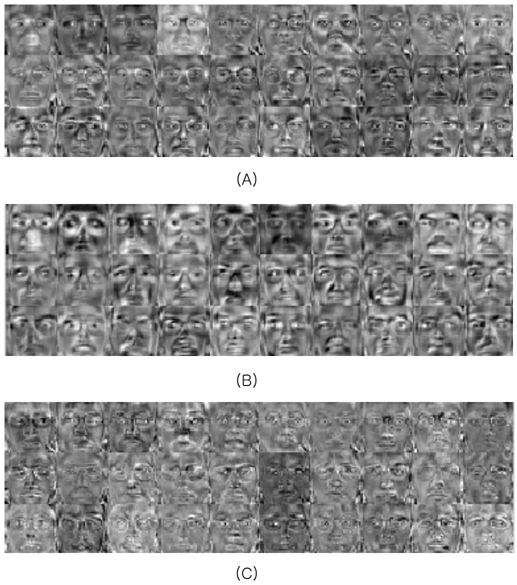


Fig. 2. Three different basis images from different method, (A) traditional PCA and LDA, (B) LDA with pose subgroup PCA, (C) LDA with illumination subgroup PCA.

eyes and nose than Fig. 2 (B), because images of the illumination set have changed overall intensity by lighting with a fixed face angle.

3. COMBINING OF SEVERAL LDA WITH SUBGROUP PCA METHODS

As a result of each LDA with subgroup PCA, we have N classifiers and their scores. The i th similarity score between probe and model images are calculated by Euclidean distance as

$$S_i(\alpha,\beta) = \left\| y_i(\alpha) - y_i(\beta) \right\|, \qquad (5)$$

where S_i is the matching score, α and β are probe and model image respectively. Such similarity scores are combined by the weighted sum, that is

$$S_f(\alpha,\beta) = \sum_{i=1}^{N} w_i S_i(\alpha,\beta), \qquad (6)$$

where the weighting factor w_i , $1 = \sum_{i=1}^{N} w_i$, were heuristically chosen according to the performance of each classifier. In this paper, to determine the weighting factor, the class discriminability of each classifier is computed from the training data set and the factor is then proportional to the discriminability value. Fig. 1 (C) shows the proposed scheme of the composed LDA with subgroup PCAs and its merging procedure.

4. EXPERIMENTAL RESULTS AND DISCUSSION

4.1. Training and Test Database Organization

The database for performance evaluation is a union of five different databases as shown at Table 1. The ten face images of each individual were selected from the origin databases because some images are not coincided with the experimental purpose such as occluded face images of Purdue database. In this paper, Purdue[5] and PIE database[6] are considered as the illumination data set, and others, ALTKOM and XM2VTS database considered as the pose data set. The training set and test set consist of even persons and odd persons of each database, respectively. All images of BANCA database are utilized as a part of the test set to evaluate time elapse and outdoor environmental changes. Namely, the whole database consists of 3,600 images of 360 individuals with 10 images of each person, as shown at Fig. 3 as examples. The images were normalized by manual eye positions, resized to 32×32 pixels, and cropped to exclude the background changes.

4.2. Experimental Results and Discussion

Table 1. Database organization for experiments

DB	# face image	Training set	Test set	Variation
Purdue	920	460	460	Illumination and Expression variation
PIE	680	340	340	Illumination variation
Altkom	800	400	400	Pose and Illumination variation
Xm2vts	800	400	400	Pose variation
Banca	400		400	Illumination and Time variation

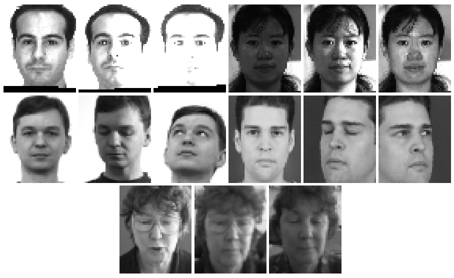


Fig. 3. Example images of used databases, from upper left, Purdue, PIE, Altkom, XM2VTS, Banca, respectively

As a measure of retrieval performance, we use ANMRR (Average Normalized Modified Retrieval Rate) specified in [7]. ANMRR is 0 when all ground truth images are ranked on top, and it is 1 when all ground images are ranked out of a defined threshold.

We evaluated six different methods for the various tests, and their descriptions are following:

- EXP1: (PCA and LDA) Learning individuals with all training sets.
- EXP2: (PCA and LDA) Learning individuals with only pose training sets; ALTKOM, XM2VTS.
- EXP3: (PCA and LDA) Learning individuals with only illumination training sets; Purdue, PIE.
- EXP4: (LDA with Pose Subgroup PCA) PCA was learned with only pose training sets, but LDA was learned with all training sets.
- EXP5: (LDA with Illumination Subgroup PCA) PCA was learned with only illumination training sets, but LDA was learned with all training sets.
- EXP6: (Proposed algorithm) Merging two similarity scores, EXP4 and EXP5, by using the weighted sum.

Fig. 4 shows overall performances as a function of the number of features for the total test data. We can see that EXP4, LDA with pose subgroup PCA is better than others except EXP6. And LDA with Subgroup PCA methods, EXP4 and EXP5 generally show better performance than EXP2 and EXP3 which are oriented to the specific data set. There is no doubt that the latter methods can learn own variations of the training data effectively, but they are overfitted so that other variations out of the training data can not be controlled. In detail, EXP2 shows the best performance for the pose test sets while the performance for the illumination test sets is the worst one, and EXP3 also shows the opposite property as

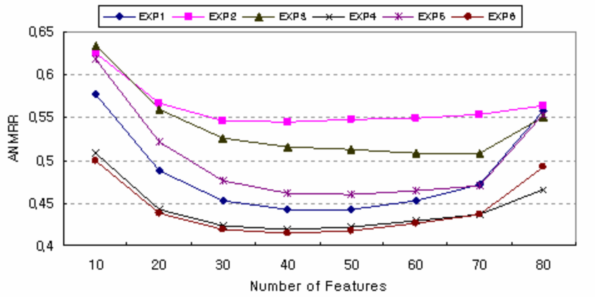
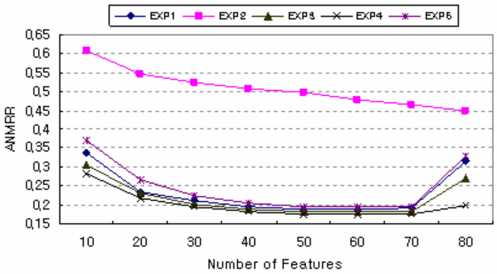
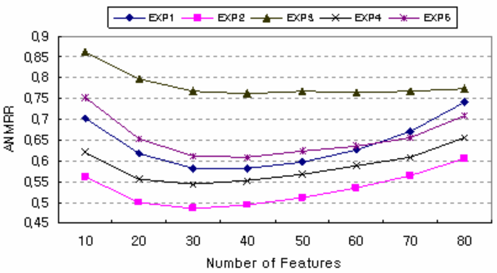


Fig. 4. Retrieving accuracies with respect to number of features in total data set



(A) Illumination Data set



(B) Pose Data set

Fig. 5. Average retrieving accuracies with respect to number of features (A) for Illumination data set, and (B) for Pose data set

shown in Fig. 5. We can overcome these serious overfitting problems by applying the proposed scheme. This result is clarified in Table 2. EXP2 learned with only pose face images shows the best performance, ALTKOM = 0.378 and XM2VTS = 0.612, for the pose based sets but also shows the worst performance, Purdue = 0.686 and PIE = 0.326, for the illumination based sets, and EXP3 show the similar inclinations. On the contrary, subgroup PCA and LDA method does not show this overfitting problem. Moreover, pose subgroup PCA and LDA shows better performance than traditional PCA and LDA. That is to say, Pose PCA subspace has more discriminability in facial image retrieval than others. The proposed merged method gave an additional improvement of the performance from 0.444 to 0.417 compared with the traditional PCA and LDA method. Considering the results of the each varying test set, we could see the similar improvements except PIE set and all the results are not

Table 2. Retrieval results for 40 feature vectors used

Unit: anmrr	Illumination based sets		Pose based sets		Outdoor	Total
	Purdue	PIE	ALTKOM	XM2VTS	BANCA	
EXP1	0.350	0.037	0.480	0.686	0.620	0.444
EXP2	0.686	0.326	0.378	0.612	0.677	0.547
EXP3	0.349	0.030	0.680	0.844	0.636	0.517
EXP4	0.322	0.042	0.440	0.664	0.591	0.420
EXP5	0.368	0.043	0.521	0.697	0.640	0.464
EXP6	0.315	0.038	0.441	0.657	0.594	0.417

the best ranks, but always higher ranks.

5. CONCLUSION

We have presented a novel approach which enables the appearance-based facial image retrieval techniques to successfully deal with different categories of face variations. The subgroups are designed to learn the local variations effectively at PCA stage with less loss of information but LDA can learn affluent discriminability with whole training sets. Moreover, the method exploits several classifiers of LDA with subgroup PCAs for the pose subgroup and illumination subgroup and they are combined together to achieve a higher recognition rate. Our research suggests one effective way to solve the overfitting problem and to improve the performance of the conventional PCA-LDA method in the scenario of face image retrieval.

6. REFERECES

[1] W. Zhao, R. Chellapa, A. Rosenfeld, and P. Phillips, “Face recognition: A literature survey,” UMD CfAR Technical Report CAR-TR-948, 2000.
[2] W. Zhao, A. Krishnaswamy, R. Chellappa, D. L. Swets, J. Weng, "Discriminant Analysis of Principal Components for Face Recognition," 3rd Int. Conference on Automatic Face and Gesture Recognition, pp. 336-341, 1998.
[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, “Eigenface vs. Fisher faces: Recognition using class specific linear projection,” IEEE Trans. PAMI, vol. 19, no. 7, pp. 711-720, Jul, 1997.
[4] X. Lu and A. K. Jain, “Resampling for Face Recognition,” 4th International Conference AVBPA 2003, Guildford, pp. 869-877, UK, Jun, 2003.
[5] A.M. Martinez and R. Benavente, “The AR Face Database,” CVC Technical Report #24, Jun, 1998.
[6] T. Sim, S. Baker and M. Bsat, “The CMU Pose, Illumination, and Expression (PIE) Database,” 5th International Conference on AFGR, Washington, D.C., May, 2002.
[7] B.S. Manjunath, P. Salembier, and T. Silkora, *Introduction to MPEG-7: Multimedia Content Description Interface*, John Wiley & Sons LTd. 2002.