Subspace Learning with Enriched Databases Using Symmetry

Konstantinos Papachristou, Anastasios Tefas, and Ioannis Pitas

Aristotle University of Thessaloniki
Department of Informatics
Box 451, 54124 Thessaloniki, Greece
{kpapaxristou,tefas,pitas}@aiia.csd.auth.gr

Abstract. Principal Component Analysis and Linear Discriminant Analysis are of the most known subspace learning techniques. In this paper, a way for training set enrichment is proposed in order to improve the performance of the subspace learning techniques by exploiting the a-priori knowledge that many types of data are symmetric. Experiments on artificial, facial expression recognition, face recognition and object categorization databases denote the robustness of the proposed approach.

Keywords: Subspace Learning, Data Enrichment, Symmetry, Principal Component Analysis, Linear Discriminant Analysis

1 Introduction

Everyday, a vast amount of images and videos are available from many sources, resulting in the need to handle and use this information intelligently by many systems such as robotics, multimedia retrieval and recognition (face, object, etc). This means that image processing methods are a key field in computer vision applications. Many of these methods exploit subspace learning techniques which have been employed in many computer vision and pattern recognition tasks [1, 2]. Such techniques calculate projection vectors in order to reduce the data dimensionality, maintaining the meaningful information and, thus, they can be employed for dimensionality reduction, data visualization and compression, as well as as a main preprocessing step in classification and clustering methods. Some of them are unsupervised, such as Principal Component Analysis (PCA) [3], Independent Component Analysis [4], Locality Preserving Projections [5] and Non-negative Matrix Factorization [6]. Another category of SL techniques is supervised and uses the class label information of data, e.g., Linear Discriminant Analysis (LDA) [7], Discriminant Non-negative Matrix Factorization [8], Clustering based Discriminant Analysis [9] and Subclass Discriminant Analysis [10].

The aforementioned techniques do not work well when the available samples are not truly representative of the corresponding patterns. Our aim is to propose a training set enrichment approach in order to produce more representative training sets and, therefore, to improve the performance of subspace learning techniques by adding the symmetric version of each sample. This approach is based on the fact that symmetry is a main characteristic of several data types, such as faces, objects, etc.

The remainder of this paper is organized as follow. In Section 2, the subspace learning techniques, namely PCA and LDA, are briefly described. In Section 3, the proposed approach for improving the robustness of the subspace learning techniques using the symmetric versions of images are presented. In Section 4, we present experiments conducted in order to evaluate the proposed approach. Finally, conclusions are drawn in Section 5.

2 Subspace Learning Techniques

In this section, we provide a brief review of well known subspace learning techniques Principal Component Analysis in subsection 2.1, LDA in subsection 2.2 and their combination in subsection 2.3. In the following, we will consider the set $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$ to be the sample images $\mathbf{x}_i \in \mathcal{R}^{m\times 1}$ in vectorized form, while the projection vectors are denoted by $\mathbf{w} \in \mathcal{R}^{m\times 1}$. The total number of samples in the dataset, the total number of classes and the mean vector of the entire data set are denoted by N, c and $\boldsymbol{\mu}$, respectively. The initial dimensionality of the samples is denoted by m, while the dimensionality of the projection space is denoted by m'.

2.1 Principal Component Analysis

PCA tries to find projection vectors $\mathbf{w}$ that maximize the variance of the projected samples $y_i = \mathbf{w}^T\mathbf{x}_i$, for better representation. If we define the total scatter matrix $\mathbf{S}_T$ as:

$$\mathbf{S}_T = \sum_{i=1}^{N} (\boldsymbol{x}_i - \boldsymbol{\mu})(\boldsymbol{x}_i - \boldsymbol{\mu})^T, \tag{1}$$

the objective of PCA is to find the transformation matrix $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, ...\mathbf{w}_{m'}]$ that maximizes the trace of $\mathbf{S}_T$:

$$J(\mathbf{W}) = \arg\max_{\mathbf{W}} \mathrm{tr}[\mathbf{W}^T\mathbf{S}_T\mathbf{W}]. \tag{2}$$

The solution of (2) is given by the solution of following generalized eigenvalue decomposition problem:

$$\mathbf{S}_T \cdot \mathbf{w} = \lambda \cdot \mathbf{w} \tag{3}$$

keeping the m' eigenvectors of $\mathbf{S}_T$ that correspond to the m' largest eigenvalues. We can choose m' such that the sum of the m' largest eigenvalues is more than a percentage $P\%$ of the sum of the total eigenvalues.

2.2 Linear Discriminant Analysis

LDA determines projection vectors $\mathbf{w}$ so that the classes of the samples are well discriminated. For this reason, the between-class scatter matrix:

$$\mathbf{S}_B^{LDA} = \sum_{i=1}^{c} (\boldsymbol{\mu}_i - \boldsymbol{\mu}) (\boldsymbol{\mu}_i - \boldsymbol{\mu})^T \tag{4}$$

and the within-class scatter matrix:

$$\mathbf{S}_W^{LDA} = \sum_{i=1}^{c} \sum_{k=1}^{n_i} \left(\boldsymbol{x}_k^i - \boldsymbol{\mu}_i\right) \left(\boldsymbol{x}_k^i - \boldsymbol{\mu}_i\right)^T , \tag{5}$$

are defined, where $\boldsymbol{x}_k^i$ is the k-th sample in the class i and, $\boldsymbol{\mu}_i$, n_i are the mean vector and the number of samples in class i, respectively.

The objective of LDA is to find the transformation matrix $\mathbf{W}$ that maximizes the ratio of the trace of the between-class scatter to the trace of the within-class scatter matrix:

$$J(\mathbf{W}) = \arg\max_{\mathbf{W}} \frac{\mathrm{tr}[\mathbf{W}^T \mathbf{S}_B^{LDA} \mathbf{W}]}{\mathrm{tr}[\mathbf{W}^T \mathbf{S}_W^{LDA} \mathbf{W}]}. \tag{6}$$

The solution of (6) is approximated [19] by the following generalized eigenvalue decomposition problem:

$$\mathbf{S}_B^{LDA} \cdot \mathbf{w} = \lambda \cdot \mathbf{S}_W^{LDA} \cdot \mathbf{w}, \tag{7}$$

by keeping the m' eigenvectors that correspond to the m' largest eigenvalues. Because $\mathbf{S}_B^{LDA}$ is the sum of c matrices in (Equation 4) of rank one or less and only $c-1$ of these are independent, the maximum number of nonzero eigenvalues is equal to $c - 1$. Consequently, the upper bound on m' is $c - 1$.

2.3 Principal Component Analysis plus Linear Discriminant Analysis

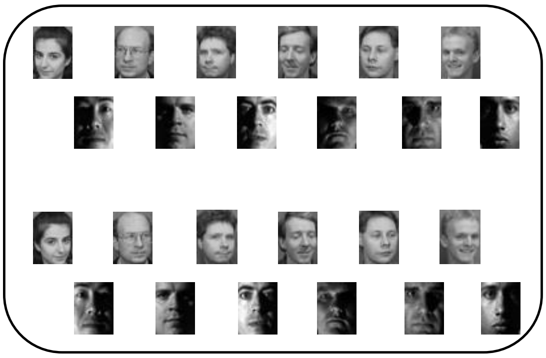
LDA is very prone to the "small sample size" problem [1]. This problem occurs when the number of samples is smaller than the dimensionality of the samples. As a result, the matrix $\mathbf{S}_W^{LDA}$ may become singular, and solving the generalized eigenvalue decomposition problem (7) may result to irregular discriminant projection vectors.

In order to overcome the above problem, an alternative technique has been proposed [11], which consists of two steps. In the first step, the samples are projected to a subspace of dimensionality lower than $N - l$ using PCA, where l denotes the number of classes for LDA technique, so that $\mathbf{S}_W^{LDA}$ become non-singular. In the second step, the matrices $\mathbf{S}_B^{LDA}$ and $\mathbf{S}_W^{LDA}$ are calculated by using the data representations in the PCA space. Finally, LDA is applied for the determination of regular projection vectors.

3 Proposed Approach

The above mentioned subspace learning techniques are rather sensitive when the training set consists of a small number of samples, resulting in a bad pattern learning and generalization. For example, as illustrated in Figure 1(a), the training set of a face recognition problem may be comprised of frontal and slightly left pose face images or face images taken with a specific light position (right). This fact can lead to a poor pattern representation by applying a subspace learning technique. A possible solution to address this problem is the enrichment of the training set by adding the symmetric version of each sample based on the symmetry property of the face in order to produce a training set which will better represent a symmetric pattern. Indeed, the application of this way of database enrichment to the images of Figure 1(a) leads to forming an enriched training set which represents better the faces of persons, as shown in Figure 1(b). These image have been inverted with respect to the vertical axis. Similarly, we can apply a corresponding training set enrichment by inverting the images with respect to the horizontal axis or to any directional axis.

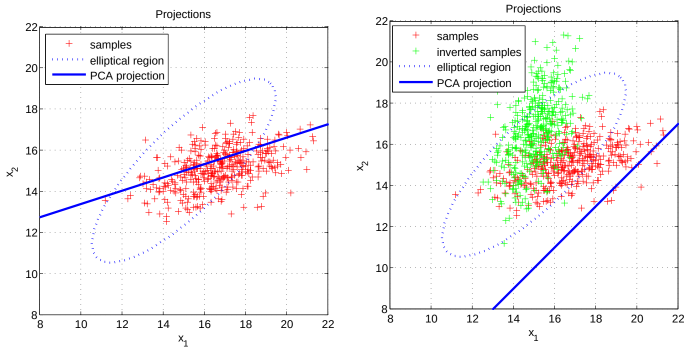
(a)



(b)

Fig. 1. Training set example consisting of (a) the original samples, and (b) both original samples and their symmetric versions.

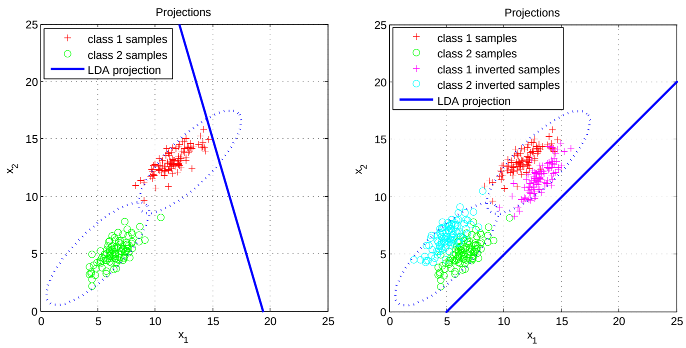
To highlight the effectiveness of the proposed training set enrichment approach in the subspace learning techniques, we designed two artificial data problems for PCA and LDA, respectively. Figure 2 illustrates the result of PCA for a symmetric artificial data problem, where the real symmetric pattern is defined by an ellipse, while the available samples are represented by crosses. As can be seen, the available samples do not correspond to a representative subset of the pattern. As a result, the PCA projection line, maximizing the samples variance, is not suitable for the real symmetric pattern. On the contrary, it is obvious that PCA results to a better projection line when the symmetric versions of samples are used.



(a) (b)

Fig. 2. PCA projection lines using (a) the original samples, and (b) both original samples and their symmetric versions.

Correspondingly, we designed an artificial two-class data problem, in which the available samples of the two classes are represented by crosses and circles, respectively. As it can be easily observed in Figure 3, LDA is able to find a projection line, which optimally separates both the available samples and the real symmetric patterns using the enriched training set compared to using the available samples only.



(a) (b)

Fig. 3. LDA projection lines using (a) the original samples, and (b) both original samples and their symmetric versions.

4 Experiments

In this section, an experimental evaluation of the proposed approach on real-databases for facial expression recognition, face recognition and object categorization is presented. We conducted two series of experiments. In the first one, the training and testing set consist of the original images of databases. In the second one, the original images and their symmetric versions were used to form the training set, while the testing set consists of the original images. In all the experiments, we applied a subspace learning technique, namely PCA, LDA and PCA+LDA, to the training set and the samples are projected into the corresponding subspace. The new dimensionality of PCA has been defined by maintaining the 99% of the total eigenvalue sum of the training set energy, while in LDA technique the new dimensionality was $c-1$, where c is the number of classes. Finally, the projected samples were classified using the Nearest Centroid (NC), and k-Nearest Neighbor (kNN) classifiers. kNN was used for $k = 1, 3, 5, 7, 9, 11$. In all classifiers, the Euclidean distance measure is adopted. The results of our experiments on facial expression recognition, face recognition and object categorization are presented in subsections 4.1, 4.2 and 4.3, respectively.

4.1 Experiments on Facial Expression Recognition

The COHN-KANADE [12] and JAFFE [13] face databases were used in our experiments for facial expression recognition. Each facial image belongs to one of the following seven facial expressions: anger, disgust, happiness, fear, sadness, surprise and neutral. The COHN-KANADE database contains 210 subjects of age between 18 and 50 years. We used 35 images of each facial expression. The JAFFE database contains 213 images depicting 10 Japanese female subjects. 3 images per subject of each facial expression were used in our experiments. All facial images were cropped to include only the subject's facial region. The cropped face images were resized to 30×40 pixels (where 30 and 40 are the columns and rows of the image, respectively). In Figure 4, a cropped facial image for all facial expressions of the COHN-KANADE and JAFFE databases is shown, respectively.

The application of LDA technique on the above databases encounters computational difficulties due to the "small sample size" problem. To estimate the recognition accuracy, we used the 5-fold cross validation procedure by dividing each database into 5 non-overlapping subsets. Each experiment included five training-test procedures (folds), where in each fold, the techniques were trained by using 4 subsets and testing was performed on the remaining subset. Recognition accuracy was measured by using the mean classification rate over all five folds. For the COHN-KANADE experiments, each subset contained 20% of the facial images for each class based on random selection. For the JAFFE database, we performed person-independent experiments: each subset contained the entire set of the facial images from 20% of the persons. Thus, the facial images of each person were either in the training or in the test set. The results obtained for the COHN-KANADE and JAFFE databases, are shown in Table 1, where the best

(h) (i) (j) (k) (l) (m) (n)

Fig. 4. A cropped image for all facial expressions of the Cohn-Kanade (first row) and JAFFE (second row) databases: (a) neutral, (b) angry, (c) disgusted, (d) feared, (e) happy, (f) sad, and (g) surprised.

results are shown in bold. As it can be seen, an improvement in the performance is observed in the majority of the cases after the enrichment with symmetric images. Thus, such an approach can be used in order to improve the performance of subspace learning techniques.

Table 1. COHN-KANADE and JAFFE 5-fold cross validation accuracy rates.

	COHN-KANADE		JAFFE	
technique	Original	Enriched	Original	Enriched
PCA	33.88	**35.10**	38.10	**40.48**
PCA+LDA	68.98	**70.61**	**51.90**	50.00

4.2 Experiments on Face Recognition

We used the ORL [14], AR [18, 16] and Extended YALE-B [15, 17] face databases in our experiments for face recognition. The ORL database contains 400 images of 40 distinct persons (10 images each). The images were captured at different times and with different variations (lighting, position). The AR database contains over 4000 color images corresponding to 70 men's and 56 women's faces. The images were taken in frontal position with different facial expressions, illumination conditions and occlusions. Each person contains 26 images capturing in two recording sessions. The Extended YALE-B database contains images of 38 persons in 9 poses and under 64 illumination conditions. The frontal cropped images were used only, in this work. All images were resized to 30×40 pixels, in our experiments. Some example facial images from the ORL, the AR and the Extended YALE-B databases are displayed in Figure 5.

For the above databases, the 20% of images per person were randomly selected for training and the remaining images were used for testing. The direct application of the LDA technique in all the databases was impossible because

(a) (b) (c)

Fig. 5. Sample images from the (a) ORL, (b) AR and (c) Extended YALE-B databases.

of the "small sample size" problem. The results obtained for the ORL, AR and Extended YALE-B databases, are illustrated in Table 2. As can be seen, in all the cases, a better regocnition accuracy is achieved when training set is enriched with the symmetric versions of the original images. Therefore, we can conclude that for symmetric data (such as a human face) the proposed way of enriching databases achieves better data representation and overcomes the poor representation using both the original images and their symmetric versions.

Table 2. ORL, AR and Extended YALE-B Accuracy Rates.

	ORL		AR		Extended YALE-B	
technique	Original	Enriched	Original	Enriched	Original	Enriched
PCA	81.88	**83.75**	27.81	**31.95**	55.06	**55.57**
PCA+LDA	80.94	**85.63**	48.86	**53.81**	81.53	**82.77**

4.3 Experiments on Object Categorization

In the experiments on object categorization we used the ETH-80 [20] database. It contains images from eight categories: apple, pear, tomato, cow, horse, dog, cup and car. For each category there are images of ten different objects. Each object has been captured by 41 different views. The images were resized to 32×32 pixels.

Table 3. ETH-80 5-fold cross validation accuracy rates.

technique	Original	Enriched
PCA	85.43	**85.67**
LDA	74.88	**81.52**
PCA+LDA	**85.00**	84.58

We evaluated the performance of the proposed techniques using the 5-fold cross validation procedure. Specifically, images of each object were either in the training set or the test set. The results are shown in Table 3. As can be seen, after the enrichment with symmetric images, an improvement in the performance is observed (PCA and LDA cases). On the other hand, when PCA is applied first, the projected samples are not symmetric in PCA space and, therefore, the symmetric versions of the samples do not affect on the performance of LDA.

5 Conclusions

Subspace Learning techniques have been a useful tool in many applications. In this paper, we proposed an enrichment approach of the training set by adding the symmetric versions of the available samples in problems where the patterns are symmetric, for example facial expression and face recognition ones. The experiments on relevant databases and artificial data highlight that a major improvement is achieved when using subspace learning combined with symmetric enrichment training sets.

Acknowledgment

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement number 316564 (IMPART) and was partially supported by the COST Action IC1106. This publication reflects only the authors views. The European Union is not liable for any use that may be made of the information contained therein.

References

1. K. Fukunaga: Introduction to Statistical Pattern Recognition, 2nd ed. Academic Press Professional, 1990.
2. A. Jain, R. Duin, and J. Mao: Statistical Pattern Recognition: A Review. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 4–37, 2000.
3. I. Jolliffe: Principal Component Analysis, 2nd ed. Springer, 2002.
4. T.-W. Lee: Independent Component Analysis: Theory and Applications. Kluwer Academic Publishers, 1998.
5. X. He and P. Niyogi: Locality preserving projections. In Advances in Neural Information Processing Systems 16, pp. 153-160, 2003.
6. D. Lee and H. Seung: Learning the parts of objects by non-negative matrix factorization. Nature, vol. 401, no. 6755, pp. 788-791, 1999.
7. R. A. Fisher: The use of multiple measurements in taxonomic problems. Annals of Eugenics, vol. 7, no. 7, pp. 179-188, 1936.
8. S. Zafeiriou, A. Tefas, I. Buciu, and I. Pitas: Exploiting discriminant information in non-negative matrix factorization with application to frontal face verification. IEEE Transactions on Neural Networks, vol. 17, no. 3, pp. 683-695, 2006.
9. X.-W. Chen and T. Huang: Facial expression recognition: a clustering-based approach. Pattern Recognition Letters, vol. 24, no. 9–10, pp. 1295-1302, 2003.

10. M. Zhu and A. Martínez: Subclass discriminant analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 8, pp. 1274-1286, 2006.
11. D. Swets and J. Weng: Using discriminant eigenfeatures for image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 18, no. 8, pp. 831-836, 1996.
12. T. Kanade, Y. Tian, and J. Cohn: Comprehensive database for facial expression analysis. In Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition. IEEE Computer Society, pp. 46-53, 2000.
13. M. Lyons, S. Akamatsu, M. Kamachi, and J. Gyoba: Coding facial expressions with Gabor wavelets. In Proceedings of the 3rd International Conference on Face and Gesture Recognition. IEEE Computer Society, pp. 200-205, 1998.
14. F. Samaria and A. Harter: Parameterisation of a stochastic model for human face identification. In Proceedings of 2nd IEEE Workshop on Applications of Computer Vision. IEEE Computer Society, pp. 138-142, 1994.
15. A. Georghiades, P. Belhumeur, and D. Kriegman: From few to many: Illumination cone models for face recognition under variable lighting and pose. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 643-660, 2001.
16. A. Martínez and A. Kak: PCA versus LDA. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pp. 228-233, 2001.
17. K.-C. Lee, J. Ho, and D. Kriegman: Acquiring linear subspaces for face recognition under variable lighting. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 5, pp. 684-698, 2005.
18. A. Martínez and R. Benavente. The AR face database. CVC Technical Report, vol. 24, 1998.
19. H. Wang, S. Yan, D. Xu, X. Tang, and T. Huang: Trace ratio vs. ratio trace for dimensionality reduction. In IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-8, 2007.
20. B. Leibe and B. Schiele: Analyzing Appearance and Contour Based Methods for Object Categorization. In IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society, pp. 409-415, 2003.