2006년도 대한전자공학회 하계종합학술대회 제29권 제1호

Clustering을 결합한 PCA와 LDA 기반 얼굴 인식

곽련화*, 김표재**, 장형진***, 최진영****
서울대학교 공과대학 전기· 컴퓨터 공학부

Face Recognition Based on PCA and LDA Combining Clustering

Lianhua Guo*, Pyo Jae Kim**, Hyung Jin Chang***, Jin Young Choi****
School of Electrical Engineering and Computer Science,
Seoul National University
E-mail : {*lhguo, **pjkim, ***hjchang, ****jychoi}@neuro.snu.ac.kr

Abstract

In this paper, we propose an efficient algorithm based on PCA and LDA combining K-means clustering method, which has better accuracy of face recognition than Eigenface and Fisherface. In this algorithm, PCA is firstly used to reduce the dimensionality of original face image. Secondly, a truncated face image data are sub-clustered by K-means clustering method based on Euclidean distances, and all small subclusters are labeled in sequence. Then LDA method project data into low dimension feature space and group data easier to classify. Finally we use nearest neighborhood method to determine the label of test data. To show the recognition accuracy of the proposed algorithm, we performed several simulations using the Yale and ORL (Olivetti Research Laboratory) database. Simulation results show that proposed method achieves better performance in recognition accuracy.

I. Introduction

In recent years, face recognition has become a very active research area due to increasing demands of our modern lives. Numerous face recognition systems have been proposed[3][4], among which, some systems are based on well known techniques, PCA[5] and LDA[1]. Eigenface[6] is based on PCA algorithm. LDA method can be used not only for classification, but also for dimensionality reduction. The most well known technique, the Fisherface[2], combine the techniques of PCA with LDA. Fisherface is robust to illumination and facial expression and has lower error rate than Eigenface[2]. In this paper, we propose a modified method of the Fisherface to increase accuracy of face recognition. We utilize K-means clustering method in each class data after perform PCA algorithm. As a result, data are grouped into small subclasses which benefit the following LDA algorithm. LDA algorithm project data to subspace, which make data grouped more tightly and more easily to classify. After LDA, we use the nearest neighborhood method to classify the well grouped and truncated data.

The Yale data database and ORL database are the famous and commonly used database to benchmark the performance of face recognition system. We also test the proposed algorithm on these two databases. Simulation results show that the proposed method has higher accuracy than Fisherface.

II. Proposed Method

In face recognition applications, original input face image data usually has a high dimension. So PCA can be used effectively to extract most discriminating features having small dimension than original. Using these features, LDA finds an optimal linear projection to low-dimensional feature space in which we can separate each class data.

However after implementing PCA algorithm, data still has great overlaps with other class data. These overlaps make LDA work poorly. To solve this problem, we use K-means clustering method before LDA. In detail, we set same subcluster number in each class and then subcluster data of each class using K-mean clustering algorithm as shown in Fig 1. Through this process, data are grouped into many small clusters in which data is more tightly gathered as before. These closely associated subclasses make it easier to find an optimal space to classify. Then we label newly all small clusters in sequence. Using new subclass labels and LDA, we find a projection from data space to feature one. From this projection, we get a low dimensional and most important feature data. Finally, we use the nearest neighborhood method to classify test data based on Euclidean distance. The whole flow chart of proposed algorithm is shown in Fig 2.

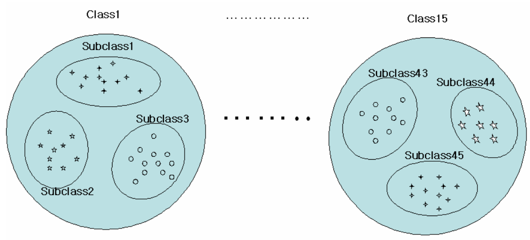

Fig 1. Sub-clustering process using K-means clustering method

Ⅲ. Simulation

We perform simulation using the proposed algorithm in matlab, and test the face recognition performance on two databases: 1) The Yale database; 2) The Olivetti Research Laboratory (ORL) database.

The Yale database contains 165 frontal face images covering 15 individuals taken under 11 different conditions. Each individual has different facial expressions, illumination conditions and small occlusion. The ORL database is composed of 400 images of 40 persons. Each person consists of 10 different facial views that represent various expressions, small occlusion (by glasses), different scale and orientations.

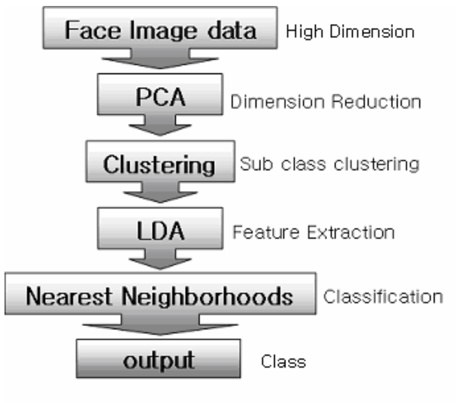

Fig 2. Block diagram of proposed algorithm

First, we subsample the large image into 30x40 size, then use PCA method to reduce the dimension of data. Second, we normalize reduced data. Third, we use K-means method to subcluster data in each class into a number of subclass. Fourth, we use LDA to project the data with the subclass label to the feature space. Finally, we use Nearest Neighborhood method to classify test data. Because the database is limited, we adopt Leave-One-Out method to test algorithms.

To test the proposed algorithm on Yale database, we reduce the dimension of face image to 35 using PCA and reduce to 13 by LDA method. Face recognition accuracy is as good as 99.4%. Two parameters about data dimension reduction are very important to the results of recognition. Eigenface method's final reduced dimension is 30, and the final accuracy is 84.85%. Fisherface method's final reduced dimension is 13, and the final accuracy is 98.12%. Compared with Eigenface and Fisherface, the proposed method has better results and has small data dimension.

Table1. Result on Yale database

Yale database		
Method	**Reduced space**	**Best accuracy (%)**
Eigenface	**30**	**84.85**
Fisherface	**13**	**98.12**
Proposed method (subclass # = 3)	**14**	**99.4**

Then we apply the proposed algorithm on ORL database. When we set the reduced dimension parameter of PCA to 38, and set the second reduced dimension parameter of LDA as 22, the final accuracy is 99.5%. The dimension of Eigenface method's final reduced space is 30, and the accuracy is 97.5%. The final dimension of Fisher space is 22, and the accuracy is 99.25%. Compared with Eigenface and Fisherface, the proposed method also has better accuracy on the ORL database.

Table2. Result on ORL database

ORL database		
Method	**Reduced space**	**Best accuracy (%)**
Eigenface	**30**	**97.5**
Fisherface	**22**	**99.25**
Proposed method (subclass # = 3)	**22**	**99.5**

Simulation results about two databases are shown in Table1 and Table2 at each. From the results, the proposed method has better performance than Eigenface and Fisherface.

IV. Conclusion

In this paper, we propose the method based on PCA and LDA combining the K-means algorithm. After using PCA algorithm, the dimension of the data is reduced. However the overlaps between classes still exist, which will adversely affect LDA algorithm to find an optimal projection space. To eliminate the shortage, we use K-means algorithm as the subclustering method and label subclasses of each class orderly. This process contributes LDA method to seek an optimal projection subspace. Although Fisherface, based on PCA and LDA, is an effective and robust algorithm. The proposed method combining K-means method improves the accuracy of Fisherface. From the simulation results on two databases, we can conclude that the accuracy of proposed algorithm has better accuracy than Eigenface and Fisherface.

For the future work, we will use other clustering methods to improve the performance of proposed algorithm.

Reference

[1] Richard O. Duda, Peter E. Hart, David G. Stork, *Pattern Classification*, Jone Wiley & Sons, INC, 2001

[2] Peter N. Belhumeur, Joao P. Hespanha, and David J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no.7, pp.711-720, Jun. 2000

[3] R. Chellappa, C. Wilson, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," Proc. IEEE, vol. 83, no. 5, pp. 705- 740, 1995.

[4] A. Samal and P. Iyengar, "Automatic Recognition and Analysis of Human Faces and Facial Expressions: A Survey," Pattern Recognition, vol. 25, pp. 65-77, 1992.

[5] I. T. Jollife, *Principal Component Analysis*, Springer-Verlag, 1986

[6] M. A. Turk and A. P. Pentland, "Face Recognition using Eigenfaces," In Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pages 586-591, 1991.