Face Recognition Using Complete Fuzzy LDA

Wankou Yang[1], Hui Yan[1], Jianguo Wang[1,2], Jingyu Yang[1]
[1] *School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China*
[2]*Tangshan College, Network & Education Center, Tangshan, 063000, China*
wankou_yang@yhoo.com.cn, {betty,wangjianguo}@njust.edu.cn, yangjy@mail.njust.edu.cn

Abstract

In this paper, we propose a novel method for feature extraction and recognition, namely, Complete Fuzzy LDA (CFLDA). CFLDA combines the complete LDA and fuzzy set theory. CFLDA redefines the fuzzy between-class scatter matrix and fuzzy within-class scatter matrix that make fully of the distribution of sample and simultaneously extract the irregular discriminative information and regular discriminative information. Experiments on the Yale and FERET face databases show that CFLDA can work well and surpass Fuzzy Fisherface.

1. Introduction

Feature extraction (or dimensionality reduction) is an important research topic in computer vision and pattern recognition fields, since (1) the curse of high dimensionality is usually a major cause of limitations of many practical technologies; (2) the large quantities of features may even degrade the performances of the classifiers when the size of the training set is small compared to the number of features [1]. In the past several decades, many feature extraction methods have been proposed, in which the most well-known ones are PCA and LDA [2-7]. Furthermore, J.Yang et al proposed Complete LDA [7,8] by considering the discriminative information, regular and irregular.

Recently, Fuzzy Fisherface was proposed for feature extraction and face recognition [9]. Fuzzy Fisherface computed fuzzy within-class scatter matrix and between-class scatter matrix by incorporating class membership of the binary labeled faces (patterns). Although it was proved to be effective, Fuzzy Fisherface did not completely incorporate the class membership into the definition of between-class and within-class scatter matrices and ignored the discriminative information in the null space of fuzzy within-class scatter matrix. So we propose Complete Fuzzy LDA (CFLDA) for feature extraction and face recognition.

2. Complete Fuzzy LDA (CFLDA)

Before presenting our method, some remarks on the Fuzzy Fisherface should be made first.

1) The sample distribution information is not completely used in the definitions of fuzzy between-class and within-class scatter matrices.

2) In PCA transformed space, the fuzzy within-class scatter matrix still might be singular.

3) The null space of the fuzzy within-class scatter matrix SF_w contains discriminative information for classification. This kind of information is ignored in the Fuzzy Fisherface method.

In view of the above three remarks, we will try to take into account these factors and manage to find a better method for feature extraction and recognition. In this section, we propose a new Complete Fuzzy LDA (CFLDA) algorithm, which makes fully of the distribution of samples and considers the discriminative information in the null space of fuzzy within-class scatter matrix. Samples distribution information is represented by fuzzy membership degree corresponding to every class.

2.1 Fuzzy K-Nearest Neighbor (FKNN)

How can we completely represent the distribution of these samples and improve classification performance through extracting discriminative information from these samples? Obviously, fuzzy set theory is a good choice.

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

In our method, fuzzy membership degree and each class center are gained through FKNN [10] algorithm. With FKNN algorithm, the computations of the membership degree can be realized through a sequence of steps:

Step1: Compute the Euclidean distance matrix between pairs of feature vectors in training set.

Step2: Set diagonal elements of this Euclidean distance matrix to infinity.

Step3: Sort the distance matrix (treat each of its columns separately) in an ascending order. Collect the corresponding class labels of the patterns located in the closest neighborhood of the pattern under consideration (as we are concerned with ‘ k ’ neighbors, this returns a list of ‘ k ’ integers).

Step4: Compute the membership degree to class ‘ i ’ for jth pattern using the expression proposed in the literature [10]

$$u_{ij} = \begin{cases} 0.51 + 0.49 \times (n_{ij}/k) & \text{if i = the same as the jth label of the pattern} \\ 0.49 \times (n_{ij}/k) & \text{if i } \neq \text{ the same as the jth label of the pattern} \end{cases} \quad (1)$$

In the above expression n_{ij} stands for the number of the neighbors of the jth data (pattern) that belong to the ith class. As usual, u_{ij} satisfies two obvious properties:

$$\sum_{i=1}^{c} u_{ij} = 1$$

$$0 < \sum_{j=1}^{N} u_{ij} < N \quad (2)$$

Taking into account the fuzzy membership degree, the mean vector of each class is

$$\tilde{m}_i = \frac{\sum_{j=1}^{N} u_{ij}^{p} x_j}{\sum_{j=1}^{M} u_{ij}^{p}} \quad (3)$$

where p is a constant which controls the influence of fuzzy membership degree.

Therefore, the class center matrix m and the fuzzy membership matrix U can be achieved with the result of FKNN.

$$U = \left[u_{ij} \right], \; i = 1, 2 \cdots c, \; j = 1, 2, \cdots N \quad (4)$$

$$m = \left[\tilde{m}_i \right], \; i = 1, 2 \cdots c \quad (5)$$

2.2 The Idea of Complete Fuzzy LDA

The key step of Complete Fuzzy LDA is how to incorporate the contribution of each training sample into the redefine of scatter matrices. With the conception of fuzzy set theory, every sample can be classified into multi classes under fuzzy membership degree, which is different to binary classification problem. The results of the FKNN are used in the computations of the statistical properties of the patterns. Then, the membership degree of each sample (contribution to each class) should be considered and the corresponding fuzzy within-class scatter matrix and fuzzy between-class scatter matrix can be redefined as follow:

$$FS_w = \sum_{i=1}^{c} \sum_{x_j \in w_i} u_{ij}{}^{p} (x_j - \tilde{m}_i)(x_j - \tilde{m}_i)^T \quad (6)$$

$$FS_b = \sum_{i=1}^{c} \sum_{j=1}^{N} u_{ij}^{p} (\tilde{m}_i - m)\tilde{m}_i - m)^T \quad (7)$$

where p is same as p in Eq.(3), m is the mean of all samples. So, all scatter matrices with fuzzy set theory is redefined and the contribution of each sample is incorporated.

Our optimal fuzzy projection W_{CFLDA} follows the expression:

$$W_{CFLDA} = \arg\max_{W} \frac{\left| W^T FS_b W \right|}{\left| W^T FS_w W \right|} \quad (8)$$

It is difficult to directly calculate W_{CFLDA} because that FS_w is often singular. According to the idea of Complete LDA [7], we firstly reduce the dimension into the range space of the total scatter matrix; in the PCA transformed space, we split the fuzzy within-class scatter matrix FS_w into its null space $\Phi_w^{\perp} = span\left\{ v_{q+1}, \cdots, v_m \right\}$ and its orthogonal complementary $\Phi_w = span\left\{ v_1, \cdots, v_q \right\}$, where $v_1, \cdots, v_m$ are orthonormal eigenvectors of FS_w, and the first q ones are corresponding to positive eigenvalues. Then, we will extract the discriminative vectors in these subspaces.

Our proposed method makes fully of the distribution information of samples, derives the irregular discriminative vectors from the null space of FS_w and regular discriminative vectors from the complementary space of FS_w , so it is called Complete Fuzzy LDA (CFLDA).

3. Experiments

3.1 Experiments Using FERET Database

The FERET face database is a result of the FERET program, which was sponsored by the US Department of Defense through the DARPA Program [12,13]. It has become a standard database for testing and evaluating state-of-the-art face recognition algorithms. The proposed method was tested on a subset of the FERET database. This subset includes 1,400 images of 200 individuals (each individual has seven images). This subset involves variations in facial expression, illumination, and pose. In our experiment, the facial portion of each original images was automatically cropped based on the location of the eyes and the cropped images was resized to 80×80 pixels. Some examples images of one person are shown in Fig.1.

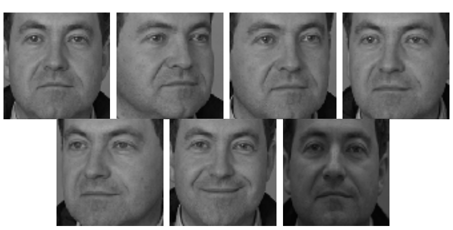

Figure 1. Images of one person in FERET

In the experiment, we use the first l ($l=4,5$) images per class for training and the remaining images for testing. For feature extraction, we used, respectively, PCA(eigenface) [3], LDA(Fisherface)[6], Fuzzy Fisherface(F-Fisherface) [9], Complete LDA(C-LDA)[7] and the proposed method. Note that PCA, Fisherface and Fuzzy Fisherface all involve a PCA phase. In this phase, we keep nearly 98 percent image energy and select the number of principal components, m , as 375 and 433. The K -nearest neighborhood parameter K can be chosen as $K=l-1$. The justification for this choice is that each sample should connect with the remaining $l-1$ samples of the same class provided that within-class samples are well clustered in the observation space.The parameter p in the complete fuzzy LDA is set as $p=2$ .Finally a nearest neighbor classifier with cosine distance is employed. The final recognition rate of each method and the corresponding dimension are given in Table2. Table 1. shows that the proposed method (CFLDA) has a better performance than others.

Table 1 Recognition Comparision On FERET

	4	5
PCA	0.4075(375)	0.3725(433)
LDA	0.4317(199)	0.3725(199)
F-Fisherface	0.4517(199)	0.3450(199)
C-LDA	0.4283(461)	0.3625(399)
Proposed	**0.5217(461)**	**0.4350(399)**

From Table 1, we can see that the proposed method (CFLDA) outperforms other methods. Why can CFLDA outperform other methods? In our opinion, First the overlapping sample's distribution information is completely incorporated in the redefinition of corresponding scatter matrices by fuzzy set theory, which is important for classification. Second, the discriminative information in the null space of fuzzy within-class scatter matrix is considered in the feature extraction.

3.2 Experiments Using Yale Database

The Yale face database contains 165 images of 15 individuals (each person providing 11 different images) under various facial expressions and lighting conditions. In our experiments, each image was manually cropped and resized to 100×80 . Figure 2 shows sample images of one person.

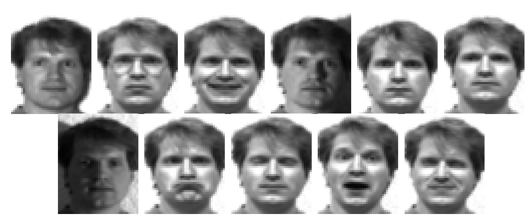

Figure 2. Eleven images of one person in Yale

The experiment was performed using the first l ($l=3,4$) images per class for training, and the remaining five images for testing. For feature extraction, we used, respectively, PCA(eigenface), LDA(Fisherface), Complete LDA(C-LDA), Fuzzy Fisherface(F-Fisherface) and the proposed method. In the PCA phase of PCA, LDA and Fuzzy Fishface, we keep nearly 98 percent image energy and select the number of principal components, m , as 34 and 43. The K -nearest neighborhood parameter K in Fuzzy Fisherface and Complete Fuzzy LDA can be chosen as $K = l - 1$. The parameter p in the Complete Fuzzy LDA is set as $p = 40$. Finally, the nearest neighbor (NN) classifier with cosine distance is employed for classification. The maximal recognition rate of each method and the corresponding dimensions are given in Table2. Table 2 shows that the proposed method (CFLDA) has a better performance than others.

Table 2 Recognition Comparision On Yale

	3	4
PCA	0.8417(34)	0.8762(43)
LDA	0.8167(14)	0.8467(14)
F-Fisherface	0.8167(14)	0.8286(14)
C-LDA	0.8750(15)	0.9048(16)
Proposed	**0.8917(15)**	**0.9143(16)**

4. Conclusions

We have proposed a Complete Fuzzy LDA method for feature extraction and face recognition by completely incorporating the distribution of the samples and considering the discriminative information in the null space of the fuzzy within-class scatter matrix. By virtue of Fuzzy K-Nearest neighbor, we can get the class membership of the binary labeled faces. This in turn allows us to compute fuzzy within-class scatter matrix and fuzzy between-class scatter matrix forming the core portion of the original LDA classifier. By doing this, we could reduce the sensitivity of the method to substantial variants between face images caused by large pose, expression or illumination variations. Further, we consider the discriminative information in the null space of fuzzy within-class scatter matrix, which is important for classification. Experimental results show that our proposed method is effective and outperforms Fuzzy Fisherface.

Acknowledgements

This project is supported by NSFC of China (Grant Nos: 60632050, 60473039), by Hi-Tech Research and Development Program of China (Grant No. 2006AA01Z119) and by the Graduates' Research Innovation Program of Higher Education of Jiangsu Province (Grant No.CX07B_118z). The authors would like to thank Dr. zhong Jin for his help.

References

[1] A.K.Jain, B.Chandrasekaran. Dimension and Sample Size Consideration in Pattern Recognition Practice. In Handbook of Statistic. Amsterdam, North Holland. 1982.

[2] M.Kirby and L.Sirovich. Application of the KL Procedure for the Characterization of Human Faces. IEE Trans. Pattern Analysis and Machine Intelligence, 1990,12(1): 103-108.

[3] Turk M, Pentland A. Eigenfaces for recognition. Journal of Cognitive Neuroscience[J], 1991, 3(1): 71-86.

[4] K. Liu, Y.Q.Cheng, J.Y.Yang X. Liu. An Efficient Algorithm for Foley-Sammon Optimal Set of Discriment Vectors by Algebraic Methods. Journal of Patter Recognition and Artificial Intelligence. 1992, 6(5):817-829

[5] Swets D L, Weng J. Using Discriminant Eigenfeatures for Image Retrieval. IEEE Trans on Pattern Analysis and Machine Intelligence[J], 1996, 18(8): 831-836.

[6] Belhumeur V, Hespanha J, Kriegman D. Eigenfaces vs Fisherfaces: Recognition Using Class Specific Linear Projection. IEEE Trans on Pattern Analysis and Machine Intelligence[J]. 1997, 19(7):711-720.

[7] J. Yang, J.Y.Yang. Why can LDA be performed in PCA transformed space?. Pattern Recognition. 2003,36(2):563-566.

[8] J.Yang, A.F.Frangi, J-Y.yang, D.Zhang, Z.Jin. KPCA Plus LDA: A complete kernel fisher discriminant framework for feature extraction and recognition. IEEE Trans. On PAMI. 2005,27(2):230-244.

[9] Kw K C, Pedry W. Face recognition using a fuzzy fisher classifier. Pattern Recognition. 2005,38(10):1717-1732

[10] Keller J M, Gray M R, Givern J. A. A fuzzy k-nearest neighbour algorithm. IEEE Trans. Syst. Man Cybernet. 1985,15(4): 580-585

[11] J.Yang, J.Y.Yang. Optimal FLD algorithm for facial feature extraction. SPIE Proceedings of the Intelligence Robots and Computer Vision XX: Algorithms, Techniques, and Active Vision, October, Vol. 4572, 2001, 438-444

[12] P.J. Phillips, H. Moon, S.A. Rizvi, and P.J. Rauss. The FERET Evaluation Methodology for Face-Recognition Algorithms. IEEE Trans. Pattern Analysis and Machine Intelligence, 2000, 22(10):1090-1104.

[13] P.J. Phillips. The Facial Recognition Technology (FERET) Database. http://www.itl.nist.gov/iad/humanid/feret/feret_master.html, 2004.