Face Recognition Using Fuzzy Discriminant Locality Preserving Projection

Pengli Lu and Xingbin Jiang
School of Computer and Communication, Lanzhou University of Technology, 730050, Lanzhou, China

Abstract: A novel approach, namely Fuzzy Discriminant Locality Preserving Projection (FDLPP), is proposed for dimensionality reduction to improve the performance of Discriminant Locality Preserving Projection (DLPP). FDLPP which is based on Maximum Margin Criterion (MMC), pursues to maximize the difference between the locality preserving between-class scatter and locality preserving within-class scatter instead of the ratio. In FDLPP, fuzzy k-nearest is implemented to obtain correct local distribution information and the pursuit of better classification results. Blending the membership degree into the definition of the Laplacian scatter matrix acquire to fuzzy Laplacian scatter matrix. Experiments on ORL, FERET and Yale face databases show the effectiveness with the change in illumination and viewing directions of the proposed method.

Key words: Discriminant locality preserving projection, fuzzy k-nearest, dimensionality reduction, maximum margin criterion, fuzzy laplacian scatter matrix

INTRODUCTION

Dimensionality reduction has been the key technology in many fields of information processing, for instance pattern recognition and data mining. In face recognition, extracting useful identifying information in the training samples which is including a lot of redundant information is necessary before classification. To overcome the difficulty, it is usually to learn a subspace in which we can detect the reduced intrinsic dimension in the high dimensional space. This needs to map the sample points form input space to a low dimensional space by linear or nonlinear transformation. Many useful techniques, two popular linear subspace methods are PCA (Turk and Pentland, 1991) and LDA (Fukunaga, 1990), for dimensionality reduction has been developed over the recent decades.

One limitation of PCA and LDA is that they effectively see only the linear global Euclidean structure but the essential structure of data with nonlinear sub-manifold cannot be explored. A lot of manifold learning-based approaches focus on preserving the local structure, such as Locally Linear Embedding (LLE) (Roweis and Saul, 2000), Laplacian Eigenmaps (LE) (Belkin and Niyogi, 2003) attracted considerable attention over the past years. The embedding function of the linear techniques is defined everywhere in the input space rather than only defined for a set of data samples such as non-linear embedding techniques (Donkiak and Assoum, 2013). This is the main superiority of dimensionality reduction using the linear approaches over the non-linear ones.

Recently, a graph-based method which is based on Locality Preserving Projections (LPP) (He and Niyogi, 2003) technique was proposed for linear dimensionality reduction (LDR) for overcome the out-of-sample problem. As a linear unsupervised technique, LPP preserves the locality structures of data is superior to PCA. Furthermore, extensive experiments shown that it can give better performance than the supervised technique LDA in some cases. Many approaches such as Optimal Locality Preserving Projection (O-LPP) (Chen et al., 2011), DLPP/MMC (Lu and Zuo, 2011), Orthogonal Isometric Projection (O-IsoProjection) (Lu et al., 2010) and Discriminant Locality Preserving Projections (DLPP) (Yu et al., 2006) are proposed based on LPP to advance the classification performance of LPP. DLPP is proposed as an approach of maximizing the average distances of between-class and minimizing the average variances for the same class.

In this study, we improve the discriminant locality preserving projections to overcome the singularity encountered by DLPP and blend the fuzzy technology to increase the worst-case performance for each individual and a Fuzzy Discriminant Locality Preserving Projections (FDLPP) algorithm is proposed. Furthermore, maximum margin criterion (Lu et al., 2010) is introduced in FDLPP. By taking advantage of the technology of fuzzy sets, a number of studies have been carried out for fuzzy image filtering, fuzzy image segmentation and fuzzy edge detection with an ultimate objective to cope with the factor of uncertainty being inherently present in many problems of image processing and pattern recognition (Kwak and Pedrycz, 2005). In fact, the FDLPP algorithm

Corresponding Author: Xingbin Jiang, School of Computer and Communication, Lanzhou University of Technology, 730050, Lanzhou, China

4340
can achieve better estimation of class means by using the membership degree to describe the distribution of training samples. Experimental results show the effectiveness of the proposed approach.

**RELATED WORKS**

Given a training set of face images \( X = \{x_1, x_2, \ldots, x_n\} \subseteq \mathbb{R}^{m \times p} \), each face image \( x_i \) belongs to one of the \( C \) face classes \( \{X_1, X_2, \ldots, X_C\} \). The objective function of DLPP can be expressed as follows (Chen et al., 2011):

\[
A = \arg \max_{B} \sum_{c=1}^{C} \sum_{i \in c} \left( \frac{y_i' \cdot \mathbf{w}_{i}}{\left\| y_i - \overline{y}_i \right\|^2} \right)
\]

where, \( n_i \) is the number of samples in the \( i \)th class, \( y_i' \) represents the \( i \)th projected vector in the \( c \)th class, \( m_i \) and \( \overline{m}_i \) is separated to measure the mean projected vector for the \( i \)th class and 1st class, \( \mathbf{w}_{i} \) represents the elements of within-class weight matrix and \( \mathbf{w}_{i} = \exp(-\left\| x_i' - x_i' \right\|^2/\delta^2) \), \( B \) represents the elements of between-class weight matrix and \( B = \exp(-\left\| f_i - f_i \right\|^2/\delta^2) \), where \( \delta \) is an empirically determined parameter, \( x_i' \) represents the \( i \)th vector in the \( c \)th class, \( f_i \) is the mean of the \( i \)th class. Thus, the between-class weight matrix is \( B - \mathbf{B} \), the within-class weight matrix is \( W = \text{diag}[W^{(1)}, W^{(2)}, \ldots, W^{(C)}] \). It is clear that both \( B \) and \( W \) are symmetric positive semi-definite matrices.

Suppose that the mapping from \( x_i \) to \( y_i' \) is A, then, the objective Eq. 1 can be rewritten as:

\[
J(A) = A^{T}FH^{T}A
\]

where, \( L \) and \( H \) are Laplacian matrices. \( L = D \cdot W \), \( D = \text{diag}\{D_1, D_2, \ldots, D_C\} \), \( D \) is a diagonal matrix and its elements are column (or row) sum of \( W^{(i)} \); \( H = E - B \), \( E \) is a diagonal matrix and its elements are column (or row) sum of \( B \). \( F = \{f_1, f_2, \ldots, f_C\} \).

**PROPOSED FDLPP**

The results of the FKNN are used in the computations of the statistical properties of the patterns such as the mean value and scatter covariance matrices, the constructs being at heart of the Fisherface method. Taking into account the membership grades, the mean vector of each class \( m_i \) is calculated as follows:

\[
m_i = \frac{\sum_{j=1}^{n_i} u_{ij} x_j}{\sum_{j=1}^{n_i} u_{ij}}
\]

According to Eq. 3, the fuzzy between-class weight matrix \( B_{\text{between}} \) are redefined as follows:

\[
B_{\text{between}} = \exp(-\left\| m_i - \overline{m}_i \right\|^2/\delta^2)
\]

A reasonable criterion for choosing a 'good' map is to minimize the following fuzzy within-class scatter under appropriate constraints:

\[
\sum_{i=1}^{C} \sum_{j=1}^{n_i} (y_i' - y_i') \mathbf{U} \mathbf{W}_{ij} \mathbf{U}^T
\]

\[
= \sum_{i=1}^{C} \sum_{j=1}^{n_i} (y_i' - y_i') \mathbf{U} \mathbf{W}_{ij} \mathbf{U}^T
\]

\[
= \sum_{i=1}^{C} \sum_{j=1}^{n_i} y_i' x_i' \mathbf{U} \mathbf{W}_{ij} \mathbf{U}^T
\]

\[
= \mathbf{x}' \mathbf{X}_{\text{between}} \mathbf{x}' - \mathbf{x}' \mathbf{X}_{\text{between}} X_1 \mathbf{x}'
\]

where, the fuzzy within-class weight matrix \( \mathbf{W}_{\text{within}} \) is defined as follows:

\[
\mathbf{W}_{\text{within}} = \mathbf{U}^T \cdot \exp(-\left\| u_{i} x_i' - u_{j} x_j' \right\|^2/\delta^2)
\]

Then the fuzzy Laplacian matrices \( L_{\text{between}} \) and \( H_{\text{between}} \) are calculate as:

\[
L_{\text{between}} = D_{\text{between}} - W_{\text{between}}
\]

\[
H_{\text{between}} = E_{\text{between}} - B_{\text{between}}
\]

where, \( D_{\text{between}} \) is a diagonal matrix and its elements are column (or row) sum of \( W_{\text{between}} \), \( E_{\text{between}} \) is a diagonal matrix and its elements are column (or row) sum of \( B_{\text{between}} \).

Finally, by combining Eq. 6-7, the objective Eq. 2 can be redefined as:

\[
J(A) = A^{T}MH_{\text{between}} \mathbf{X} \mathbf{X}^T A
\]

From the classical Fisher criterion function, we know the larger the ratio of the between-class scatter to the within-class, the easier samples can be separated.

FDLPP requires that matrix \( \mathbf{X} \mathbf{X}^T \) be nonsingular. For many applications involving the SSS problem this matrix is singular. In fact, as long as the dimension of the sample \( M \) is greater than the number of samples \( N \), \( \mathbf{X} \mathbf{X}^T \) must be singular. Thus FDLPP cannot be applied directly. The goal of MMC for feature extraction is to seek \( k \) discriminant vectors \( a_1, a_2, \ldots, a_k \) such that the trace of locality preserving between-class scatter is maximized and the trace of locality preserving within-class scatter is minimized after training samples are projected on these vectors. According to the idea of MMC (Lu et al., 2010), the objective function of FDLPP is defined as follows:
\[ j(A) = \text{tr}(A^T(S_\alpha - \alpha S_w)A) \]  

(8)

where, \( S_\alpha = MH^TM^T \), \( S_w = aX_fX_f^T \).

The parameter \( \alpha \) is a nonnegative constant which balances the relative merits of maximizing the fuzzy between-class scatter to the minimization of the fuzzy within-class scatter.

In real-world application of such face recognition, gene expression and web document recognition, the dimension \( M \) of the vector samples is usually large, so performing FDLPP by directly solving the eigenvectors of the \( M \times M \) matrix \( S_\alpha - \alpha S_w \) is still computationally intensive. To reduce the computational demand, in this section, we present an efficient algorithm for performing FDLPP.

In the following, for the sake of simplicity, we assume that we deal with centered data. Suppose \( \beta_1, \beta_2, \ldots, \beta_m \) are \( M \) orthonormal eigenvectors of total scatter matrix \( S_\alpha \) and the first \( m \) (\( m = \text{rank} \, S_\alpha \)) ones \( \beta_1, \beta_2, \ldots, \beta_m \) are eigenvectors of \( S_\alpha \), corresponding to positive eigenvalues and \( \beta_{m+1}, \beta_{m+2}, \ldots, \beta_m \) are eigenvectors of \( S_\alpha \) corresponding to zero eigenvalues. Define the subspace \( \Phi_i = \text{span}\{\beta_1, \beta_2, \ldots, \beta_i\} \) and its orthogonal complement can be denoted by \( \Phi_i^\perp = \text{span}\{\beta_{m+1}, \beta_{m+2}, \ldots, \beta_m\} \). Obviously, \( \Phi_i^\perp \) is the null space of \( S_\alpha - S_\alpha + S_w \).

**Lemma 1:** Lu et al. (2010) Suppose \( \Phi_i^\perp \) is the null space of \( S_\alpha \), then for every \( \zeta \in \Phi_i^\perp \), we have \( \zeta^T x_i = 0, \ i = 1, 2, \ldots, N \).

**Theorem 1:** Lu et al. (2010) Suppose \( P = [\beta_1, \beta_2, \ldots, \beta_m] \) to be the matrix of all unit eigenvectors of \( S_\alpha \), corresponding to nonzero eigenvalues and \( u \in \mathbb{R}^m \) to be the eigenvector of the matrix \( P^T(S_\alpha - \alpha S_w)P \) corresponding to the eigenvalues \( \lambda \). Then \( Pu \) is the eigenvector of the matrix \( S_\alpha - \alpha S_w \) corresponding to the eigenvalues \( \lambda \).

Now, the algorithmic procedure of FDLPP is formally summarized as follows:

**Step 1:** Calculate the membership degree \( U \)

**Step 2:** Construct the fuzzy within-class weight matrix \( W_\text{fuzzy} \) and fuzzy between-class weight matrix \( B_\text{fuzzy} \) according to Eq. 3 and 5

**Step 3:** Calculate the fuzzy within-class Laplacian matrix \( L_\text{fuzzy} \) and fuzzy between-class Laplacian matrix \( H_\text{fuzzy} \) by Eq. 6-7

**Step 4:** Calculate the matrix \( P \) of all eigenvectors of \( S_\alpha \) corresponding to positive eigenvalues

**Step 5:** Solve the eigenvalues problem of \( P^T(S_\alpha - \alpha S_w)P \)

\( Pu = \lambda u, \ u_1, u_2, \ldots, u_k \) are orthonormal eigenvectors corresponding to the \( k \) largest eigenvalues

**Step 6:** The optimal projection matrix is given by \( A = PU \)

**EXPERIMENTS**

To evaluate the performance of the proposed FDLPP method, we have performed experiments on ORL, Yale and FERET database. Experiments investigate the performance of the FDLPP algorithm over the reduced dimensions and compare it with the DLPP, DLMPP and other feature extraction algorithms. Experimental results demonstrate the accuracy of FDLPP over the variance of the dimensionality of subspaces. Figure 1 shows sample images of one person in ORL, FERET and Yale databases.

**ORL database:** The ORL database (AT and T Lab Cambridge) contains 400 face images of 40 individuals which were taken at different times, varying illumination and variations in facial expression. The total number of

---

**Fig. 1:** Samples of face image in ORL, FERET and yale face databases

---

4342
Fig. 2(a-c): Accuracies on ORL database with the variance of dimension (a) 4-Trains, (b) 5-Trains and (c) 6-Trains

Fig. 3(a-c): Accuracies on FERET database with the variance of dimension (a) 3-Trains, (b) 4-Trains and (c) 5-Trains
images for each individual is 10. Each image was digitized and stored as an 112×92 pixel array whose gray levels ranged between 0 and 255. Each image in ORL database was manually cropped and resized to 64×64 in our experiments. The first 1 images (1 varies from 4-6) of each individual are selected for training while the rest images are used for testing.

**FERET database:** The FERET database (Phillips et al., 2000) contains 1400 images of 200 individuals, each individual has seven images which were taken under unstable illumination, variations in facial expression. In our experiments, each image was manually cropped and resized to 60×60. The first 1 images (1 = 3, 4, 5) of each individual are selected for training while the remaining images are used for testing.

**Yale database:** The Yale face database (Yale University, 2002) consists of 165 images from 15 individuals under various facial expressions and lighting conditions. Each individual has 11 images which was manually cropped and resized to 60×60. The first 1 images (1 = 4, 5, 6) of each individual are selected for training and the remaining images are used for testing.

**CONCLUSION**

In this study, FDLPP method is proposed for feature extraction and face recognition. The fuzzy set theory is implemented to achieve the definition of the fuzzy Laplacian scatter matrix. The effect of the outlier’s classification can be effectively reduced. The range both of the locality preserving between-class scatter $S_b$ and the
locality preserving within-class scatter $S_w$ can be derived from the discriminant vector by MNC. Extensive experiments shown the improved accuracy and the reduced sensitivity to variations between face images caused by changes in illumination are superior to other feature extraction methods.

ACKNOWLEDGMENT

Project supported by the National Natural Science Foundation of China (11361033) and the Natural Science Foundation of Gansu (1212RJZA029).

REFERENCES


