2015 International Conference on Intelligent Information Hiding and Multimedia Signal Processing

Extended Discriminant Nearest Feature Line Analysis for Feature Extraction

Yunxia Liu*, Tie Cai, Guowei Huang
School of Computer Science
Shenzhen Institute of Information Technology
Shenzhen, China
liuyx@sziit.edu.cn

***Abstract*—In this paper, a novel feature extraction algorithm, entitled Extended Discriminant Feature Line Analysis (EDFLA), is proposed. EDFLA is a Nearest Feature Line (NFL) metric based dimensionality reduction method. For small size sample problem, the existing prototype samples usually are not enough to describe the corresponding class. To extend the representation ability of the prototype sample set, a novel prototype sample set will be generated in EDFLA using the original prototype samples and NFL. EDFLA aims at minimizing the within-class scatter and maximizing the between class scatter of the novel generated prototype sample set. The experimental results on COIL20 image database and AR face database confirm the effectiveness of the proposed algorithm.**

***Keywords*-Feature line; subspace learning; feature extraction;**

I. Introduction

Dimensionality reduction methods are viewed as one of the most important steps for such fields as computer vision, image classification, and machine learning [1]. Samples in these areas are often recognized as a point in a high-dimensional space, such as gene expression profile, digital images text document videos, etc. With the development of biometric technology, more effective and more efficient dimensionality reduction approaches are required. Due to high dimensionality, many applications usually encounter a well-known problem of ”curse of dimensionality”. Dimensionality reduction has three main goals: 1) lower time and space complexities for computing; 2) low-dimensional visualizations for data analysis; 3) denoising and improving the classification or prediction performance. For linear dimensionality reduction methods, they usually map the samples the original samples to low-dimensional subspace using a linear transformation matrix. Principal Component Analysis (PCA) [2], [3] and Linear Discriminant Analysis (LDA) [4], [5] are two classical approaches for unsupervised and supervised dimension reduction, respectively.

PCA aims to seek an optimal linear transformation matrix to maximize the explained variance of the samples in the feature space. PCA perfomrs the linear unsupervised dimensionality reduction by an optimal transformation matrix from the eigenvectors of the covariance matrix of the samples. However, it generally perform badly for image classification problems due to the unsupervised nature. LDA is a linear supervised dimensionality reduction method which maximizes the between-class variance and minimizes the within-class variance jointly. LDA has two main limitations: 1) For the traditional LDA, the dimensionality of the projected subspace can be $c-1$ at most, where c is the number of classes, and 2) it assumes that each class follows a unimodal distribution, which may not always hold.

In resent years, several methods based on image matrix are introduced, such as Two-Dimensional Principal Component Analysis (2DPCA) [6] and Two-Dimensional Linear Discriminant Analysis (2DLDA) [7]. 2DPCA and 2DLDA can compute the optimal transformation matrix and extract the features in a straightforward manner based on the image matrix projection. And these matrix-based algorithms, not only reduce the computational complexity greatly, but also improve the recognition effectiveness. Many approaches based on image matrix are presented in recent years [8], [9].

Most of the above methods use Euclidean distance to measure the scatter of the samples. In recent years, some methods based on other metric are proposed. Nearest Feature Line (NFL) is a classifier proposed by Lee in 1999. Its main idea is using feature line metric to measure the distance between prototype sample and some class. Zheng et al. proposed NFL-based Nonparametric Discriminant Analysis (NFL-NDA) [10] in 2006 and Pang et al. propose a Nearest Feature Line-based Space (NFLS) [11] in 2007. NFLS aims to minimize the with-in class scatter of the prototype samples based on NFL. Uncorrelated Discriminant Nearest Feature Line Analysis (UDNFLA) [12] is proposed by Lu and Tan in 2010. UDNFLA is based on the theory of NFL with additional to extracting the uncorrelated discriminant features of the prototype samples. The experimental results of these algorithms show that the performance of these algorithms are pretty efficient. Discriminant Nearest Feature Space Analysis (DNFSA) [13] is presented by J. Lu and Y. P. Tan in 2011. NFS is applied to extract the discriminant features of prototype samples in DNFSA. Yan et al. proposed the Neighborhood Discriminant Nearest Feature Line Analysis (NDNFLA) [14] and some improve algorithms based on NFL [15].

978-1-5090-0188-0/15 $31.00 © 2015 IEEE
DOI 10.1109/IIH-MSP.2015.113

II. PRELIMINARIES

A. Nearest Feature Line

Nearest feature line is a classifier. It is first presented by Stan Z. Li and Juwei Lu. Given a training samples set, $X = \{x_n \in R^D : n = 1, 2, \cdots, N\}$, denote the class label of x_i by $l(x_i)$, the training samples sharing the same class label with x_i by $P(i)$, and the training samples with different label with x_i by $R(i)$. NFL generalizes each pair of prototype feature points belonging to the same class: $\{x_m, x_n\}$ by a linear function $L_{m,n}$, which is called the Feature Line (FL). The line $L_{m,n}$ is expressed by the span $L_{m,n} = sp(x_m, x_n)$. The query x_i is projected onto $L_{m,n}$ as a point x^i_{mn}. This projection can be computed as

$$x^i_{mn} = x_m + t(x_n - x_m) \tag{1}$$

where $t = [(x_i - x_n)(x_m - x_n)]/[(x_m - x_n)^T (x_m - x_n)]$.

The Euclidean distance between x_i and x^i_{mn} is termed as Feature Line metric. The less the Feature Line metric is, the more probability that x_i belongs to the same class as x_m and x_n. Fig. 1 shows a sample of Feature Line metric. In Fig. 1, the distance between y_p and the feature line $L_{1,2}$ equals to the distance between y_q and y_p, where y_p is the projection point of y_q to the feature line $L_{1,2}$.

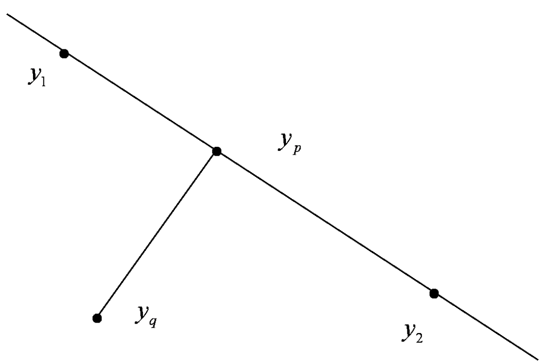

Figure 1. Feature Line metric

B. Extended Nearest Feature Line Space

In this section, a novel feature extraction algorithm called Extended Nearest Feature Line Space (ENFLS) is proposed. For many image classification problems, usually there are only a few prototype samples in one class. It is very important to get enough and proper samples to describe one class. NFL can generate new samples using FLs. In NFL, all points in the FL are viewed as the prototype samples in the corresponding class. Given a prototype samples set $X = \{x^j_i \in R^D, i = 1, \cdots, c, j = 1, \cdots, n_i\}$, where x^j_i denotes the jth sample in the ith class. For the ith, there are $C^2_{n_i}$ FLs, where $C^2_{n_i}$ denotes the combined number. For each prototype sample in the ith class, one generated samples can be got by projection to one FL in the same class. Except the original prototype sample, each prototype sample can generate $C^2_{n_i-1}$ samples. For the ith class, $n_i \times C^2_{n_i-1}$ new samples can be generated using this method. In the novel ENFLS algorithm, all generated samples are viewed as prototype samples. Then the number of the prototype sample set increases by $\sum_{i=1}^{c} n_i \times C^2_{n_i-1}$ samples. Suppose that the extended prototype sample set is $Y = \{y^j_i \in R^D, i = 1, \cdots, c, j = 1, \cdots, N_i\}$, where $N_i = n_i \times (1 + C^2_{n_i-1})$. Let $N = \sum_{i=1}^{c} N_i$. The novel prototype sample set constitutes of n_i original samples and $n_i \times C^2_{n_i-1}$ generated samples. ENFLS aims to minimize the intra-class scatter of the novel prototype sample set. Therefore the optimization function of the proposed ENFLS is as follows:

$$\min J(W) = \sum_{i=1}^{c} \sum_{j=1}^{N_i} \left\| Z^j_i - EZ_i \right\|^2, \tag{2}$$

where $Z^j_i = W^T Y^j_i$, $EZ_i = \sum_{j=1}^{N_i} Z^j_i \Big/ N_i$.

The optimization problem in Eq. (2) can be transformed to the following eigenvalue problem:

$$Sw = \lambda w, \tag{3}$$

where

$$S = \sum_{i=1}^{c} \sum_{j=1}^{N_i} ((Y^j_i - EY_i)(Y^j_i - EY_i)^T), \tag{4}$$

$$EY_i = \frac{1}{N_i} \sum_{i=1}^{N_i} Y^j_i. \tag{5}$$

III. THE PROPOSED ALGORITHM

In the last section, ENFLS algorithm extracts the features of the images using the within-class scatter based on feature line metric. For classification task, between class scatter is also very important. In this section, a novel image feature extraction algorithm called Extended Discriminant Feature Line Analysis (EDFLA) is proposed. This method also can be used for small size sample problem. Similar to ENFLS algorithm, EDFLA also uses NFL to generate some new samples. In the prototype sample set, each sample and each feature line in one class can generate one novel sample if this sample does not belong to the corresponding feature line. That is the sample and two samples which generate the feature line are not collinear. Given a prototype samples set $X = \{x^j_i \in R^D, i = 1, \cdots, c, j = 1, \cdots, n_i\}$, where x^j_i denotes the jth sample in the ith class. For the ith, there are $C^2_{n_i}$ FLs, where $C^2_{n_i}$ is the combined number. For each prototype sample in the ith class, one generated samples can be computed by projection to one FL in the same class. Except the original prototype sample, each prototype sample can generate $C^2_{n_i-1}$ novel samples. For the ith class, $n_i \times C^2_{n_i-1}$ new prototype samples can be generated using this method. In the proposed EDFLA algorithm, all the novel generated samples are viewed as prototype samples. Then the number of the prototype sample set increases by $\sum_{i=1}^{c} n_i \times C^2_{n_i-1}$

samples. Suppose that the extended prototype sample set is $Y = \{y_i^j \in R^D, i = 1, \cdots, c, j = 1, \cdots, N_i\}$, where $N_i = n_i \times (1 + C_{n_i-1}^2)$. Let $N = \sum_{i=1}^{c} N_i$. The novel prototype sample set constitutes of n_i original samples and $n_i \times C_{n_i-1}^2$ generated samples. EDFLA aims to minimize the within-class scatter and maximize the between class sactter of the novel prototype sample set of the extended prototype sample set. Therefore the optimization function of the proposed EDFLA is as follows:

$$\begin{aligned} \min J(W) = & \sum_{i=1}^{c}\sum_{j=1}^{N_i}\sum_{p=1,p\neq i}^{c} \left\| Z_i^j - EZ_p \right\|^2 \\ & - \sum_{i=1}^{c}\sum_{j=1}^{N_i} \left\| Z_i^j - EZ_i \right\|^2 \end{aligned}, \quad (6)$$

where $Z_i^j = W^T Y_i^j$, $EZ_i = \sum_{j=1}^{N_i} Z_i^j \Big/ N_i$.

$$\begin{aligned} & \sum_{i=1}^{c}\sum_{j=1}^{N_i} \left\| Z_i^j - EZ_i \right\|^2 \\ & = \sum_{i=1}^{c}\sum_{j=1}^{N_i} \mathrm{tr}((Z_i^j - EZ_i)(Z_i^j - EZ_i)^T) \\ & = \mathrm{tr} \sum_{i=1}^{c}\sum_{j=1}^{N_i} (W^T(Y_i^j - EY_i)(Y_i^j - EY_i)^T W) \\ & = \mathrm{tr} W^T \left(\sum_{i=1}^{c}\sum_{j=1}^{N_i} ((Y_i^j - EY_i)(Y_i^j - EY_i)^T) \right) W \\ & = \mathrm{tr} W^T S_w W, \end{aligned} \quad (7)$$

where

$$S_w = \sum_{i=1}^{c}\sum_{j=1}^{N_i} ((Y_i^j - EY_i)(Y_i^j - EY_i)^T), \quad (8)$$

$$EY_i = \frac{1}{N_i}\sum_{i=1}^{N_i} Y_i^j. \quad (9)$$

$$\begin{aligned} & \sum_{i=1}^{c}\sum_{j=1}^{N_i}\sum_{p=1,p\neq i}^{c} \left\| Z_i^j - EZ_p \right\|^2 \\ & = \sum_{i=1}^{c}\sum_{j=1}^{N_i}\sum_{p=1,p\neq i}^{c} \mathrm{tr}((Z_i^j - EZ_p)(Z_i^j - EZ_p)^T) \\ & = \mathrm{tr} \sum_{i=1}^{c}\sum_{j=1}^{N_i}\sum_{p=1,p\neq i}^{c} (W^T(Y_i^j - EY_p)(Y_i^j - EY_p)^T W) \\ & = \mathrm{tr} W^T \left(\sum_{i=1}^{c}\sum_{j=1}^{N_i}\sum_{p=1,p\neq i}^{c} ((Y_i^j - EY_p)(Y_i^j - EY_p)^T) \right) W \\ & = \mathrm{tr} W^T S_b W, \end{aligned} \quad (10)$$

where

$$S_b = \sum_{i=1}^{c}\sum_{j=1}^{N_i}\sum_{p=1,p\neq i}^{c} ((Y_i^j - EY_p)(Y_i^j - EY_p)^T). \quad (11)$$

Then the object of the proposed EDFLA can be transformed as follows:

$$J = S_b - S_w. \quad (12)$$

Table I
MARR OF DIFFERENT ALGORITHMS ON COIL20 DATABASE

Approaches	MARR	Feature dimension
PCA	0.6521	75
LDA	0.6986	19
LPP	0.6758	45
NFLS	0.7016	100
UDNFLA	0.7856	80
ENFLS	0.8192	18
The proposed EDFLA	0.8327	22

Then the optimization problem in Eq. (6) equals to the following eigenvalue problem:

$$(S_b - S_w)w = \lambda w. \quad (13)$$

Let $w_1, \cdots, w_d$ be the eigenvectors corresponding to the d smallest eigenvalues of the matrix $S_b - Sw$ and $W = [w_1, \cdots, w_d]$. Then W is the optimal transformation matrix for the proposed EDFLA. Denote $Z_i^j = W^T Y_i^j \in R^d$. Z_i^j is the feature extracted by EDFLA and will be used for classification instead of the original sample Y_i^j.

IV. Experimental results

To evaluate the effectiveness of the proposed ENFLS approach, we compare it with PCA, LDA, LPP, UDNFLA, NFLS and ENFLS on the COIL20 [16] database and AR face image database [17].

A. Experimental results on COIL20 database

COIL20 database is one of most well-known image databases. There are more than 1400 images of 20 classes in the COIL20 database. Each class has 72 image samples with the size of 200 × 200. All the image samples in COIL database are cropped to 32 × 32 in the following experiments to reduce the computation complexity. Firstly, PCA is performed on the prototype sample set. In this stage, 97% energy is preserved. During the experiments, 10 image samples per class are randomly selected for training and the others are for test. Nearest Feature Line classifier is used for classification. The system runs ten times in order to reduce the variation on different partitions of the database. Therefore, the Maximum of Average Recognition Rate (MARR) is applied to evaluate the performance of different algorithms. Table I shows the MARRs of different algorithms on COIL20. From the Table, we can get the proposed algorithm outperforms the other popular algorithms.

B. Experimental results on AR face database

AR face database was created by Aleix Martinez and Robert Benavente in the Computer Vision Center (CVC) at the U.A.B. It contains over 4,000 color images corresponding to 126 people's faces (70 men and 56 women). Images feature frontal view faces with different illumination

Table II
MARR OF DIFFERENT ALGORITHMS ON AR DATABASE

Approaches	MARR	Feature dimension
PCA	0.6912	70
LDA	0.7201	19
LPP	0.7036	36
NFLS	0.7281	80
UDNFLA	0.7692	80
ENFLS	0.7786	24
The proposed EDFLA	0.8583	25

conditions, facial expressions, and occlusions (sun glasses and scarf). The pictures were taken at the CVC under strictly controlled conditions. Each person participated in two sessions, separated by two weeks (14 days) time. The same pictures were taken in both sessions. In the following experiments, only nonoccluded images of AR face database are selected. Three images per person are randomly selected for training and the other images are for testing. Table II shows the MARRs of different algorithms on AR database.

V. CONCLUSION

This paper has proposed a new dimensionality reduction-based feature extraction algorithm called Extended Discriminant Feature Line Analysis (EDFLA) for feature extraction. In EDFLA, to enhance the representation ability of the prototype sample set, NFL is used to generate an extended prototype sample set. EDFLA aims to minimize the within-class scatter and maximize the between class scatter of the extended prototype sample set. In the experiments, the proposed algorithm has been performed on COIL20 database ORL face image database and compared with some popular methods, such as PCA, LDA, LPP, UDNFLA, NFLS, and ENFLS. The experimental results on COIL20 database and AR face image database confirm the effectiveness of EDFLA.

ACKNOWLEDGMENT

This work was supported in part by Guangdong Natural Science FoundationRobust Fault Detection for Discrete Time-varing systems (S2012010008964) Shenzhen Science and Technology R&D funding Basic Research Program Active Fault Tolerant Control for Netorked Control Systems with Incomplete Measurements(JCYJ20120615103057639) and Guangdong provincial natural science foundation of P.R.China (Grand No. S2013010012669).

REFERENCES

[1] J.-B. Li, S.-C. Chu, J.-S. Pan, and L. C. Jain, “Multiple viewpoints based overview for face recognition,” *Journal of Information Hiding and Multimedia Signal Processing*, vol. 3, no. 4, pp. 352–369, Oct. 2012.

[2] S. Roweis, “EM algorithms for PCA and SPCA,” in *in Advances in Neural Information Processing Systems*. MIT Press, 1998, pp. 626–632.

[3] M. Turk and A. Pentland, “Eigenfaces for recognition,” *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, Jan. 1991.

[4] A. M. Martinez and A. Kak, “PCA versus LDA,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 23, no. 2, pp. 228–233, Feb 2001.

[5] P. Belhumenur, J. Hepanha, and D. Kriegman, “Eigenfaces vs. Fisherface: Recognition using class specific linear projection,” *IEEE Transactions on Pattern Analysis Machine and Intelligence*, vol. 19, no. 7, pp. 711–720, Jul. 1997.

[6] J. Yang, D. Zhang, A. Frangi, and J. Yang, “Two-dimensional PCA: a new approach to appearance-based face representation and recognition,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, Jan. 2004.

[7] J. Yang, D. Zhang, X. Yong, and J. Yang, “Two-dimensional discriminant transform for face recognition,” *Pattern Recognition*, vol. 37, no. 7, pp. 1125–1129, Jul. 2005.

[8] L. Yan, J.-S. Pan, S.-C. Chu, and K. K. Muhammad, “Adaptively weighted sub-directional two-dimensional linear discriminant analysis for face recognition,” *Future Generation Computer Systems*, vol. 28, no. 1, pp. 232–235, Jan. 2012.

[9] L. Yan, J.-B. Li, X. Zhu, J.-S. Pan, and L. Tang, “Bilinear discriminant feature line analysis for image feature extraction,” *Electronic Letters*, vol. 51, no. 4, pp. 336–338, Feb. 2015.

[10] Y. Zheng, J. Yang, J. Yang, X. Wu, and Z. Jin, “Nearest neighbour line nonparametric discriminant analysis for feature extraction,” *Electron. Lett*, vol. 42, no. 12, pp. 679–680, Jun. 2006.

[11] Y. Pang, Y. Yuan, and X. Li, “Nearest neighbour line nonparametric discriminant analysis for feature extraction,” *Electron. Lett*, vol. 42, no. 12, pp. 679–680, Jun. 2006.

[12] J. Lu and Y. P. Tan, “Uncorrelated discriminant nearest feature line analysis for face recognition,” *IEEE Signal Processing Letters*, vol. 17, no. 2, pp. 185–188, Feb. 2010.

[13] J. Lu and Y. Tan, “Nearest feature space analysis for classification,” *IEEE Signal Processing Letters*, vol. 18, no. 1, pp. 55–58, Jan. 2011.

[14] L. Yan, W. Zheng, S. Chu, and J. Roddick, “Neighborhood discriminant nearest feature line analysis for face recognition,” *Journal of Internet Technology*, vol. 14, no. 1, pp. 344–347, Jan. 2013.

[15] L. Yan, C. Wang, S.-C. Chu, and J.-S. Pan, “Discriminant analysis based on nearest feature line,” in *Intelligent Science and Intelligent Data Engineering*, ser. LNCS, Nanjing, 2013, pp. 356–363.

[16] S. A. Nene, S. K. Nayar, and H. Murase, “Columbia object image library (coil-20),” in *Technical Report CUCS-005-96*, 1996.

[17] A. M. Martinez and R. Benavente, “The AR face database,” in *CVC Technical Report 24*, 1998.