Palmprint Recognition Based on Local Fisher Discriminant Analysis

Jinyu Guo
College of Information Engineering, Shenyang University of Chemical Technology, Shenyang, China
Email: shandong401@sina.com

Haibin Chen and Yuan Li
College of Information Engineering, Shenyang University of Chemical Technology, Shenyang, China
Email: haibin8749@foxmail.com, li-yuan@mail.tsinghua.edu.cn

***Abstract*—A new palmprint recognition method based on local Fisher discriminant analysis(LFDA) is proposed. In order to solve the singularity of the eigenvalue equation matrix in small-size-sample cases such as image recognition, image down-sample is first used to reduce the palmprint space dimensionality. The LFDA is applied to extract the low projection vectors. Then the training images and test images are projected onto the projection vectors to get the local palmprint feature vectors. Finally, the cosine distance between two feature vectors is calculated to match palmprint. The new algorithm is tested in PolyU plmprint database. The results show that compared with principal component analysis (PCA), Fisher discriminant analysis (FDA), independent component analysis (ICA), and kernel principal component analysis (KPCA), the recognition rate of the new algorithm is the highest which is 98.95%, and the recognition time is 0.031s, so it meets the real-time system specification.**

***Index Terms*—principal component analysis (PCA), Fisher discriminant analysis(FDA), independent component analysis(ICA); kernel principal component analysis(KPCA); local Fisher discriminant analysis(LFDA)**

I. INTRODUCTION

Palmprint recognition is a new research direction in biometrics recognition field. It has become one hot research of biometrics recognition technology. There are three popular feature extraction approaches, most of which are based on structural features, frequency field features and subspace features. Structural features such as feature points and lines are difficult to be extracted and represented, but also they are not strong enough for palmprint recognition. Frequency field features ignore the texture information of image, result in larger instability subject to the variations of illumination conditions, and the image features are still high-dimensional space in line with the image dimensions. For subspace features, the works that appear in the literature include principal component analysis (PCA)[1], independent component analysis(ICA)[2,3,4], kernel principal component analysis(KPCA)[5], and Fisher discriminant analysis(FDA) [6]. The subspace feature extraction method has strong representation, low computation, easy to implement and good separation, so they are widely used in many fields, such as face recognition, palmprint recognition and so on.

PCA decorrelates the input data using second-order statistics. In a task such as palmprint recognition, in which important information may be contained in the high-order relationships among pixels, it seems reasonable to use the method sensitive to these high-order statistics. ICA is one such method. PCA and ICA are powerful linear technique, and they can not extract the nonlinear feature vectors effectively. KPCA describes the nonlinear correlations between pixels. The feature vectors of PCA, ICA and KPCA extracting are the best describe characteristics, and not the best classification characteristics. FDA seeks the projection directions that are advantageous for discrimination. In other words, the class separation is maximized in these directions. FDA tends to give undesired results if samples in some class form several separate clusters, i.e., multimodal. In this paper, a new palmprint recognition method called local Fisher discirminant analysis(LFDA) is proposed. LFDA takes local structure of the data into account, so the multimodal data can be embedded appropriately, and the better results can be obtained.

The organization of this paper is as follows: Section 2 describes the LFDA algorithm. In Section 3, a palmprint recognition algorithm based on LFDA is presented. Section 4 presents the experimental results using the PolyU palmprint database. Section 5 offers our conclusions.

II. LOCAL FISHER DISCRIMINANT ANALYSIS

FDA is a popular method for linear supervised dimensionality reduction. FDA seeks for an embedding transformation such that the between-class scatter is maximized and the within-class scatter is minimized. It is essentially achieved through generalized eigenvalue

Manuscript received January 24, 2013; revised July 12, 2013; accepted July 25, 2013.

Corresponding author: li-yuan@mail.tsinghua.edu.cn(Yuan LI)

doi:10.4304/jsw.9.2.287-292

analysis of the between-class scatter matrix S_b against the within-class scatter matrix S_w

$$S_b p = \lambda S_w p \tag{1}$$

wher λ and p are generalized eigenvalue and eigenvector, respectively. The definitions of between-class and within-class scatter matrices can be found in the literatures[6,7]. The drawback of FDA method lies in the fact that it ignores the localities of data class while focusing on the global features only. To overcome the deficiency of regular FDA method, LFDA has been proposed with both between-class separation and within-class local structure preservation simultaneously [8,9,10].

Let $x_i \in \Re^d (i = 1,2,\cdots,n)$ be d-dimensional samples and $y_i \in (1,2,\cdots,c)$ be associated class labels, where n is the number of samples and c is the number of classes. Let n_l be the number of samples in class l:

$$\sum_{l=1}^{c} n_l = n \tag{2}$$

In LFDA algorithm, two weighting matrices $\tilde{W}_b$ and $\tilde{W}_w$ are introduced into the local between-class and within-class scatter matrices as follows:

$$\tilde{S}_b = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \tilde{W}_b^{(i,j)} (x_i - x_j)(x_i - x_j)^{\mathrm{T}} \tag{3}$$

and

$$\tilde{S}_w = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \tilde{W}_w^{(i,j)} (x_i - x_j)(x_i - x_j)^{\mathrm{T}} \tag{4}$$

where the weighting matrices are defined as

$$\tilde{W}_b^{(i,j)} = \begin{cases} A_{i,j}(\frac{1}{n} - \frac{1}{n_l}) & if \quad y_i = y_j \in l \\ \frac{1}{n} & if \quad y_i \neq y_j \end{cases} \tag{5}$$

$$\tilde{W}_w^{(i,j)} = \begin{cases} \frac{A_{i,j}}{n_k} & if \quad y_i = y_j = l \\ 0 & if \quad y_i \neq y_j \end{cases} \tag{6}$$

The affinity matrix A has been used to characterize the closeness between all pairs of samples and its (i,j)th element $A_{i,j}$ is defined to quantify how far apart the data pair x_i and x_j are. One standard way of setting the $A_{i,j}$ value is as follows

$$A_{i,j} = \exp(-(\|x_i - x_j\|^2)/\sigma_i\sigma_j) \tag{7}$$

LFDA tries to keep in-class data pairs close(since $\tilde{W}_w^{(i,j)}$ is positive and $\tilde{W}_b^{(i,j)}$ is negative if $y_i = y_j = l$) and between-class data pairs apart(since $\tilde{W}_b^{(i,j)}$ is positive if $y_i \neq y_j$).

Similarly, the generalized eigenvalue analysis can be conducted between the local between-class and within-class scatter matrices

$$\tilde{S}_b \tilde{p} = \tilde{\lambda} \tilde{S}_w \tilde{p} \tag{8}$$

where the eigenvector $\tilde{p}$ corresponds to the local Fisher discriminant direction and the eigenvalue $\tilde{\lambda}$ indicates the local separation ratio between different classes. Then we weight each generalized eignvector by the square root of its associated generalized eigenvalue, that is,

$$T_{LFDA} = (\sqrt{\tilde{\lambda}_1}\, \tilde{p}_1 \Big| \sqrt{\tilde{\lambda}_2}\, \tilde{p}_2 \Big| \cdots \Big| \sqrt{\tilde{\lambda}_k}\, \tilde{p}_k) \tag{9}$$

where $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_k$. An efficient implementation of LFDA is summarized as a pseudo code in Fig.1.

III. PALMPRINT RECOGNITION BASED ON LFDA

Based on the above LFDA algorithm, a new palmprint recognition approach is proposed. The algorithm can be described as follows:

Step 1: Sample image. We preprocess the palmprint image to extract region of interest(ROI). Each ROI in the training set is divided into smaller blocks and each block contains $P \times Q$ pixels. We can obtain sampled sub-images by taking one pixel from the same position in each block[11]. In fact, P and Q is the sampling interval in height and width respectively. If the size of the palmprint image is $M \times N$ pixels, these sampled sub-images can be represented mathematically as

$$A_{ij}(i,j) = A((i-1)\times P + 1, (j-1)\times Q + 1) \tag{10}$$

where $1 \leq i \leq M / P$, $1 \leq j \leq N / Q$. Fig.2 shows the sampling result from the palmprint image on the left using Eq.(10). If the number of palmprint samples n is larger than $M / P \times N / Q$, the singular problem of the within-class scatter matrix can be solved. Transform the sampled sub-image into a one-dimensional row vector x_i.

Step 2: Select good projection vectors. All the training images consist of sample matrix X as the input of LFDA. Compute the eigenvectors and eigenvalues for the generalized eigenvector problem. Let $\tilde{p}_0$, $\tilde{p}_1$,⋯, $\tilde{p}_{k-1}$ be the solutions of (8), ordered according to their eigenvalues, $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \cdots \geq \tilde{\lambda}_k$. In LFDA, the value of k can be controlled by setting a threshold θ as follows:

$$\sum_{i=1}^{k} \tilde{\lambda}_i / \sum_{i=1}^{n} \tilde{\lambda}_i \geq \theta \tag{11}$$

Step 3: Project. After projecting each palmprint image $x_i (i = 1,2,\cdots n)$ onto the vectors T_{LFDA} , we obtain the feature vector $f_i = T_{LFDA}{}^{\mathrm{T}} \cdot x_i$.

Step 4: Match feature vectors. The feature vectors f_i and f_j are extracted from two palmprint images. Feature matching is to calculate the cosine distance $D = \frac{|f_i f_j|}{\|f_i\| \|f_j\|}$. If D is larger than a predetermined threshold D_h, the two palmprint images are from the same person. Otherwise they are not.

IV. Experimental Results and Analysis

A. The PolyU Palmprint Database

The PolyU palmprint database (http://www4.comp.polyu.edu.hk/~biometrics/) is one of the common largest image databases in palmprint field. The database contains 600 grayscale images corresponding to 100 different palms. Six samples from each of these palms were collected in two sessions, where 3 samples were captured in the first session and the other 3 samples in the second session. The average interval between the first and the second collection was two months. The resolution of all the original palmprint images is 284×384.

B. Comparison and Analysis of Recognition Results

Four images of each palmprint class are chosen randomly as training samples and the remaining two are used as testing samples. By using the preprocessing approach [4], palmprints are orientated and ROI, whose size is 128×128 pixels, is cropped to represent the whole palmprint. ROI is divided into smaller blocks and each block contains 8×8 pixels. We can obtain 16×16 sampled image by taking one pixel from the same position in each block. We reshape the sampled ROI into 1×256 vector. The training palmprint images have 400 samples, so the data matrix $\boldsymbol{X}$ is 400×256. The LFDA method is applied to $\boldsymbol{X}$ for feature extraction. In LFDA, we select the parameter θ=[0.90,0.95,0.99].

The measurement index for the accuracy of biometrics recognition technology is recognition rate (RR). The RR is mainly described by two kinds of error rates: false reject rate (FRR) and false accept rate (FAR). These two error rates are calculated by the following formula[12]:

$$\mathrm{FRR} = \frac{NFR}{NAA} \times 100\% \tag{12}$$

$$\mathrm{FAR} = \frac{NFA}{NIA} \times 100\% \tag{13}$$

where NAA and NIA are respectively the times of genuine users and imposter users trying, and NFR and NFA are respectively the times of false rejection and false accept. The formula of RR derived from FAR and FRR is as follows:

$$\mathrm{RR} = (1 - \frac{NFR + NFA}{NAA + NIA}) \times 100\% \tag{14}$$

In order to test the performance of the proposed algorithm, each test sample is matched to all the training samples. In different dimensions of palmprint space, PCA, FDA, ICA, KPCA and the proposed algorithm are respectively applied to palmprint recognition. The recognition results are shown in Table 1. From Table 1, we conclude the conclusions as following:

1) The recognition rate of FDA is higher than PCA. The feature vectors of PCA extracting are the best describe characteristics. However, FDA extracts the best classification information.
2) Compared with PCA and FDA, the recognition rate of ICA is higher than them. PCA and FDA decorrelate the input data using second-order statistics, while ICA is sensitive to high-order statistics which contain important information for recognition.
3) The recognition rate of KPCA is higher than PCA, FDA and ICA. The mainly reason is the palmprint database has illumination variations, so the nonlinear feature extraction method is better than the linear method.
4) Compared with PCA, FDA, ICA and KPCA, the recognition rate of LFDA is the highest. The best recognition rate is 98.95%. The local structure within the multimodal is retained so that the classification performance is significantly enhanced.

C. Comparison and Analysis of Running Speeds

Table 2 shows the feature extraction time, feature matching time and recognition time for different algorithms (CPU: Pentium Dual-core 2.93GHz RAM: 1.96GB) on the PolyU palmprint database under the Matlab7.10 platform. We can learn that

1) The feature extraction time of FDA is the shortest. PCA, ICA, and KPCA need to calculate the high-dimensional covariance matrix, resulting in longer feature extraction time.
2) ICA use PCA as a preprocessor to reduce computational complexity, so the feature extraction time is longer than PCA.
3) The feature extraction time of KPCA is longer than PCA and ICA. When the test sample is projected onto the vectors, the kernel matrix is first calculated, resulting in longer feature extraction time.
4) The feature extraction time of LFDA is longer than FDA. In LFDA, two weighting matrices are introduced into the local between-class and within-class scatter matrices, so that the calculation

process makes the feature extraction time longer.

5) The feature matching time for different algorithms has connection with the feature dimension and the value.

6) Among the five algorithms, the classification performance of LFDA is the best. The recognition time for LFDA algorithm is 0.031s. Therefore the algorithm is fast and effective.

V. CONCLUSIONS

A new palmprint recognition method is proposed based on LFDA. The local multimodal can be effectively retained by the distance-based weighting matrices while the class separation is simultaneously maximized by the generalized eigenvalue decomposition. The experimental results of PolyU palmprint database show that the proposed algorithm is superior to PCA, FDA, ICA, and KPCA. The best recognition rate of the LFDA is 98.95%. The recognition time is 0.031s. The proposed method meets the real-time system specification.

ACKNOWLEDGMENT

The authors wish to thank the anonymous referees for their constructive and useful comments during the development stage of the paper. This work was supported in part by a grant from education department research project of Liaoning province (No.L2012139), Liaoning province doctoral start funds (No.20131089), and national natural science foundation of China (No.61034006, No.61174119).

REFERENCES

[1] G..M.Lu, D.Zhang, and K.Q.Wang, “Palmprint recognition using eigenpalms features,” *Pattern Recognition Letters*, vol.24, pp.1463–1467, June 2003.

[2] T.Conn1e, A.T.B.Jin, M.G.K.Ong, and D.N.C.Ling, “An automated palmprint recognition system,” *Image and Vision Computing*, vol.23, pp.501–515, May 2005.

[3] L.Shang, D.S.Huang, J.X.Du, and C.H.Zeng, “Palmprint recognition using FastICA algorithm and radial basis probabilistic neural network,” *Neurocomputing*, vol.69, pp.1782–1786, August 2006.

[4] J.Y.Guo, and W.Q.Yuan, “Palmprint recognition based on independent component analysis,” *Opto-Electronic Engineering*, vol.35, pp.136–139, March 2008.

[5] M. Ekinci, and M.Aykut, “Palmprint recognition by applying wavelet based kernel PCA,” *Journal of Computer Science and Technology*, vol.23, pp.851–861, May 2008.

[6] X.Q.Wu, D.Zhang, and K.Q.Wang, “ Fisherpalms based palmprint recognition,” *Pattern Recognition Letters*, vol.24, pp.2829–2838, August 2003.

[7] J.Liu, D.Y.Xue, J.J.Cui, and X.Jia, “Palm-dorsa vein recognition based on kernel principal component analysis and Fisher linear discriminant,” Journal of Computational Information Systems, vol.8, pp.1545–1552, April 2012.

[8] J.Yu, “Localized Fisher discriminant analysis based complex chemical process monitoring,” *AIChE*, vol.57, pp.1817–1828, July 2011.

[9] M.Sugiyama, “Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis,” *Machine Learning Research*, vol.8, pp.1027–1061, May 2007.

[10] J.Yu, “Nonlinear bioprocess monitoring using multiway kernel localized Fisher discriminant analysis,” *Industrial & Engineering Chemistry Research*, vol.50, pp.3390–3402, February 2011.

[11] H.T.Yin, P.Fu, and S.W.Meng, “Sampled FLDA for face recognition with single training image per person,” *Neurocomputing*, vol.69, pp.2443–2445, October 2006.

[12] J.Y.Guo, L.H.Gu, Y.Q.Liu, Y. Li, and J.Zeng, “Palmprint recognition based on kernel locality preserving projection,” *3rd International Congress on Image and Signal Processing*, pp.1909–1913, October 2010.

Jinyu Guo received the Ph.D. degree in electrician theory and new technology from Shenyang University of Technology, Shenyang, China, in 2010. Now she’s an associate professor in Shenyang University of Chemical Technology. She has interest on the following topics: biometrics recognition algorithm, fault diagnosis algorithm, and their applications.

Haibin Chen received the B.S. degree in electronic and information engineering from Dalian University, Dalian, China, in 2009. Now he is a postgraduate in signal and information processing from Shenyang University of Chemical Technology. He has interest on the following topics: fault diagnosis algorithm and application.

Yuan Li received the Ph.D. degree in detection technology and automation equipment from Northeast University, Shenyang, China, in 2004. Now she’s a professor in Shenyang University of Chemical Technology. She has interest on the following topics: fault diagnosis algorithm and application.

Input: Labeled samples $\{(x_i, y_i) | x_i \in \Re^d, y_i \in \{1,2,\cdots,c\}\}_{i=1}^{n}$

Dimensionality of embedding space k($1 \le k \le d$)

Output: $d \times k$ transformation matrix T_{LFDA}

$\tilde{S}_b \leftarrow 0_{d\times d}$

$\tilde{S}_w \leftarrow 0_{d\times d}$

for $l = 1,2,\cdots,c$ % Compute scatter matrices in a classwise manner

$\quad \{x_i\}_{i=1}^{n} \leftarrow \{x_j\}_{j:y=l}$

$\quad$ for $i = 1,2,\cdots,n_l$ % Determine local scaling

$\quad\quad x_i^{(7)} \leftarrow$ 7th nearest neighbor of x_i among $\{x_j\}_{j=1}^{n}$

$\quad\quad \sigma_i \leftarrow \| x_i - x_i^{(7)} \|$

$\quad$ end

$\quad$ for $i, j = 1,2,\cdots,n_l$ % Define affinity matrix

$\quad\quad A_{i,j} \leftarrow \exp(-\| x_i - x_j \|^2 / (\sigma_i \sigma_j))$

$\quad$ end

$\quad X \leftarrow (x_1 | x_2 | \cdots | x_n)$

$\quad G \leftarrow X diag(A1_n) X^{\mathrm{T}} - XAX^{\mathrm{T}}$

$\quad \tilde{S}_b \leftarrow \tilde{S}_b + G/n + (1 - n_l/n) XX^{\mathrm{T}} + X1_n(X1_n)^{\mathrm{T}}/n;$

$\quad \tilde{S}_w \leftarrow \tilde{S}_w + G/n_l$

end

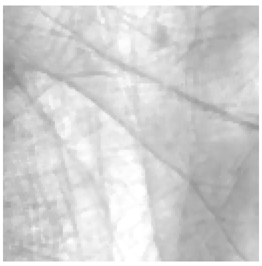
$\tilde{S}_b \leftarrow \tilde{S}_b - X1_n(X1_n)^{\mathrm{T}}/n - \tilde{S}_w$

$\{\lambda_i, p_i\}_{i=1}^{k} \leftarrow$ generalized eigenvalues and normalized eigenvectors of

$\quad \tilde{S}_b \tilde{p} = \tilde{\lambda} \tilde{S}_w \tilde{p}$ % $\tilde{\lambda}_1 \ge \tilde{\lambda}_2 \ge \cdots \ge \tilde{\lambda}_d$

$T_{LFDA} = (\sqrt{\tilde{\lambda}_1}\, \tilde{p}_1 | \sqrt{\tilde{\lambda}_2}\, \tilde{p}_2 | \cdots | \sqrt{\tilde{\lambda}_k}\, \tilde{p}_k)$

Figure 1: Efficient implementation of LFDA. $0_{d\times d}$ denotes $d\times d$ matrix with zeros, 1_n denotes the n_l-dimensional vector with ones, and diag(A 1_n) denotes the diagonal matrix with diagonal elements A 1_n.



(a) ROI of the palmprint

(b) palmprint image sampling, where P=Q=8

Figure 2. Example of palmprint image sampling.

TABLE I.
THE RECOGNITION RATE(%) FOR DIFFERENT ALGORITHMS WITH DIFFERENT FEATURE NUMBER

Algorithms	θ=0.85		θ=0.90		θ=0.95	
	Feature number	*RR*	*Feature number*	*RR*	*Feature number*	*RR*
PCA	39	91.02	59	91.68	97	91.51
FDA	18	91.45	24	91.95	35	92.66
ICA	39	94.52	59	94.11	97	94.33
KPCA	32	98.27	42	98.27	63	98.27
LFDA	32	98.89	42	98.27	63	98.27

TABLE II.
FEATURE EXTRACTION TIME(S) / FEATURE MATCHING TIME (S)/ RECOGNITION TIME(S) FOR DIFFERENT ALGORITHMS

Algorithms	Feature dimension	Feature extraction time	Feature matching time	Recognition time
PCA	97	0.026	0.007	0.033
FDA	35	0.005	0.006	0.011
ICA	97	0.030	0.007	0.037
KPCA	63	0.084	0.010	0.094
LFDA	63	0.025	0.006	0.031