Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/081L38

Improved LDA and LVQ for Face Recognition

Aijia Ouyang[1], *Kenli Li*[1,*], *Xu Zhou*[1], *Yuming Xu*[1], *Guangxue Yue*[2] *and Lizhi Tan*[1]

[1] College of Information Science and Engineering, Hunan University, Changsha, Hunan 410082, China
[2] College of Mathematics and Information Engineering, Jiaxing University, Jiaxing, Zhejiang 314001, China

Received: 14 May. 2013, Revised: 8 Sep. 2013, Accepted: 9 Sep. 2013
Published online: 1 Apr. 2014

Abstract: We present a hybrid face recognition algorithm which is based on the linear discriminant analysis (LDA) improved by a fusion technique and learning vector quantization (LVQ) in the paper. Firstly, the improved LDA is utilized to reduce the sample vector dimension, and then the LVQ classifier is used to recognize human faces. We perform intensive set of simulation experiments and results show that the algorithm not only reduces the computation time of the entire algorithm, but also improves the algorithm classification. More accurately, the algorithm obtains higher classification capability for the standard databases such as ORL, Yale and AR.

Keywords: Linear discriminant analysis, learning vector quantization, face recognition, feature extraction, dimension reduction.

1 Introduction

How to solve the face recognition has become an extremely hot research title over the past few decades and a lot of famous face recognition technologies have been designed [1] with the development of artificial intelligent technology. It can be applied to image object extraction [2] and image inpainting [3]. Among all the methods, the appearance-based method is one of the well-learned and the most successful technologies [4].

In fact, face recognition is an interesting and difficult research topic in fields such as biometrics, pattern recognition, expert system and machine vision, because it has been broadly applied to commerce and law enforcement. How to implement dimensionality reduction is very important in the processing of face recognition. The widely employed descending dimension methods can be classified into supervised methods like linear discriminant analysis (LDA) [5] and unsupervised methods like principal component analysis (PCA) [6].

The principal component analysis catches the maximum variance of data, principal component analysis is the most popular method because it is the most excellent in minimizing reconstruction error. However, limited by its nature, principal component analysis is not suitable for solving the classification problems. In detail, there is not any information class utilized in calculating the principal components.

PCA is unfit for solving classification task because it does not think about the information of class label, however, LDA can find a low-dimensional subspace by means of maximizing Fisher criterion [7]. Theoretical and experimental study demonstrated that LDA has competitive advantage over PCA for solving the problem of face recognition. LDA is very popular for the most researchers and it has been broadly applied to image processing and pattern recognition because of its effectiveness. However, LDA can not be applied directly because the within-class scatter matrix (Mw) gets singular when the dimension of sample space is much larger than the number of samples. That is to say, LDA will encounter an alleged small sample size (SSS) problem under normal circumstances. And what is more, for multi-class problems, LDA usually encounters a failure to look for fine projecting direction because it places too much emphasis on the class pairs by larger distance in the original sample space [8]. It is termed class separation problem (CSP) [9] which is originated from the fact that LDA attempts to maximize the mean distance between the centers of different classes [7] as showed by the definition of between-class scatter matrix (Mb). Many variants of LDA were proposed and showed better performance in the literatures in order to treat the above problems.

In the recent years, another linear discriminant analysis algorithm which is based on the Generalized-Singular-Value-Decomposition (GSVD), that

* Corresponding author e-mail: lkl@hnu.edu.cn

is, GSVD-based LDA (GSVD-LDA) [10], is designed to solve the small sample size problem. The GSVD-LDA acquires some discriminant vectors (DV) of itself in the whole information space because it adopts the entire scatter matrix (S_e) to replace M_w. Nevertheless, in case the information samples are not nonlinearly reliant, it maintains in lots of employments containing face images, the keys of the GSVD-LDA means will be consistent with those of the null space based technique. Then, the GSVD-LDA way will lose some valuable discriminant data too [11].

A complete linear discriminant analysis (CLDA) algorithm was presented by Yang and Yang [12] in order to solve the SSS problem of LDA. The CLDA uses the dual space of M_w, that is to say, the null space and range space of M_w, to get the DVs. Consequently, the CLDA technique may get more DVs in contrast to the other ways, thus it may extract much more valuable discriminant data.

The fitness equation of the basic linear discriminant analysis belongs to a ratio-trace-problem (RTP), namely, an approximate trace-ratio-goal (TRG). Through the TRG, the singularity problem of linear discriminant analysis can be resolved because the trace of M_w is larger than 0 all the time though M_w is singular. Nevertheless, some iterative algorithms have been developed to look for the global optimum solution (GOS) because the RTP has not a GOS OF closed-form. In the literature [13], neighborhood MAXMIN projection (NMP) by employing trace ratio objective (TRO) was addressed for solving the supervised dimensional reduction. One iterative method which can achieve the global optimum solution instead of local optimum solution was proposed. Another technique on a basis of the Newton-Raphson method released by Jia [14] who presented an efficient algorithm and fast method for solving the trace ratio objective with global optimal solution. The above techniques based on the trace ratio belongs to iterative methods, they are efficient but their time complexities are very high.

Recently, how to use the incremental studying to solve the problems of image processing has become a hot research title, therefore, a lot of incremental linear discriminant analysis (ILDA) studying algorithms have been addressed. An incremental LDA method was presented for solving the face images problem of online by Pang et al [15]. Nevertheless, they purely proposed a kind of technique which can update the scatter matrices (SM) and it will waste a lot of time, space and power; furthermore, Pang's method failed in solving the problem of SSS of linear discriminant analysis. On the basis of the GSVD-LDA method and the updating method of fast singular value decomposition [16], one incremental supervised studying means which is termed GSVD-ILDA was addressed by Zhao et al. [17]. The GSVD-ILDA may incrementally study an adaptive subspace rather than recalculating the GSVD-LDA once more; nevertheless, the way to GSVD-ILDA's calculative reduction is produced because the employment of approximate trick. In other words, the reduction of the computation will be very limited if we would like to get right singular value decomposition updating results. The technical term of the sufficient spanning set approximation (SSSA) was presented by Kim et al. [18] in updating the projected data matrix (PDM) and the between-class SM (BSM), and the entire SM (ESM). It is semblable to the GSVD-ILDA; nevertheless, there is an existing gap in the technique's performance between the answers of incremental and batch linear discriminant analysis, and because Kim's approach is not an accurate way too.

Artificial neural network [19,20,21,22,23] is an important method for face recognition, so we will emphasize the learning vector quantization (LVQ) below because LVQ integrates the advantages of the simplicity of unsupervised methods with the accuracy of supervised techniques.

LVQ, the original version was presented by Kohonen [24], is a popular means for classification. Learning vector quantization is applied in various realistic problems and engineering optimization field, they include data mining and medical image disposition, such as proteomics, language recognition, data classification, and fault-diagnosis in engineering technology. Learning vector quantization processes are very easy to perform and visually distinct. The data classification is on the basis of comparing to a good number of termed prototype vectors and the likeness is continually surveyed by Euclidean distance in feature space. Determining original models is the first thing and then interpreting it directly when they stand for classic data in the identical space. It is in the comparison with adaptive weights in support vector machines (SVM) or feedforward artificial neural networks (FANN) which permit to no instant interpretation. Among the most attractive features of learning vector quantization, naturally, where it may be employed to solve the problems of multi-class.

One thing to be noted here is that the several modifications and important improvements on Kohonen's original learning vector quantization procedure have been suggested. Their goals are to acquire better approximate values of Bayes optimal decision boundaries, rapider or more stable convergence, or the combination of more agile metrics [25,26,27].

The remainder part in our article is arranged as follows. At first, we review briefly the linear discriminant analysis method and the learning vector quantization method in Section 2. In Section 3, we address our new operation of the improved linear discriminant analysis and LVQ (ILDALVQ) method. Results of the computational experiments are presented and discussed in Section 4. Conclusion of this study is made in Section 5 in the end.

2 Standard algorithms

In this section, classic linear discriminant analysis way is stated. Then, in section 2.2, learning vector quantization schemes are presented.

2.1 statement of LDA

Assume that a data matrix $D = [d_1, d_2, \cdots, d_n] \in R^{m\times n}$, where $d_i \in R^m$, for $i = 1, \cdots, n$, d_i is the ith data point, the aim of linear discriminant analysis is to look for a transformation matrix $G^{m\times l}$ which can transform d_i into a vector from the m-dimensional space (MDS) to the l-dimensional space (LDS) as follows:

$$d_i \in R^m \rightarrow G^T d_i \in R^l \tag{1}$$

Given the data matrix D is grouped as $D = [D_1, \cdots, D_k]$,where $D_i \in R^{m\times n}$ consists of the n_i data points from the ith class and $\sum_{i=1}^{k} n_i = n$. Make N_i be the set of indices column which pertains to the ith class, i.e., d_j, for $j \in N_i$, pertains to the ith column. In linear discriminant analysis, three matrices, termed entire scatter matrices (TSM),between-class (BC) and within-class (WC), the formulations of them are given as follows [28]:

$$\begin{aligned} M_w &= \sum_{i=1}^{k} \sum_{j\in N_j} (d_j - c^{(i)})(d_j - c^{(i)})^T \\ M_b &= \sum_{i=1}^{k} n_i (c^{(i)} - c)(c^{(i)} - c)^T \\ M_t &= \sum_{j=1}^{n} (d_j - c)(d_j - c)^T = M_b + M_w \end{aligned} \tag{2}$$

we can see that the centroid $c^{(i)}$ of the ith class is denoted as $c^{(i)} = (D_i e^{(i)})/n_i$; such that, $c^{(i)} = (1, 1, \cdots, 1)^T \in R^{n_i}$, the global centroid c is denoted as $c = (De)/n$, where $c^{(i)} = (1, 1, \cdots, 1)^T \in R^n$.

Three definitions of the matrices are made respectively:

$$\begin{aligned} H_w &= [D_1 - c^{(1)}(e^{(1)})^T, \cdots, D_k - c^{(k)}(e^{(k)})^T] \in R^{m\times n} \\ H_b &= [\sqrt{n_1}(c^{(1)} - c), \cdots, \sqrt{n_k}(c^{(k)} - c)] \in R^{m\times k} \\ H_t &= D - ce^T \in R^{m\times n} \end{aligned} \tag{3}$$

After that, the above scatter matrices (SM) in (2) can also be transformed as follows:

$$\begin{aligned} M_w &= H_w H_w^T \\ M_b &= H_b H_b^T \\ M_t &= H_t H_t^T = [H_w, H_b][H_w, H_b]^T \end{aligned} \tag{4}$$

It adopts from the properties of matrix trace into this:

$$\begin{aligned} trace(M_w) &= \sum_{i=1}^{k} \sum_{j\in N_i} ||d_j - c^{(1)}||_2^2 \\ trace(M_b) &= \sum_{i=1}^{k} n_i ||c^{(i)} - c||_2^2 \end{aligned} \tag{5}$$

Therefore, the trace of M_w surveys the closeness between each data point d_i to its class centroid (CC), and the trace of M_b surveys the distances between each class centroid to the global centroid.

The three scatter matrices in the descending dimension space transformed by G, marked as M_w^L, M_b^L and M_t^L, they can be expressed as follows:

$$M_w^L = G^T M_w G, \;\; M_b^L = G^T M_b G, \;\; M_t^L = G^T M_t G \tag{6}$$

The termed Classical Fisher Criteria is expressed as follows:

$$G = arg \max_G \{trace((M_w^L)^{(-1)} M_b^L\} \tag{7}$$

Solving method on the fitness function in (7) may be gained through solving the following question of generalized eigenvalue:

$$M_b \chi = \lambda M_w \chi, \;\; \lambda \neq 0 \tag{8}$$

Whose eigenvectors which correspond to the $k-1$ largest eigenvalues form columns of G. We can see intuitively that $rank(M_b) \leq k-1$; consequently, the reduced dimension (RD) by the standard linear discriminant analysis is at most $k-1$.

2.2 The LVQ algorithm

The LVQ neural network may be expressed as a three-layer one. The fist layer is an input one. The fist layer has many neurons like pattern variables. The second layer is intermediate layer. The interlayer includes more than one prototype neuron for each class. The prototype neurons and corresponding weighted links between the prototype neurons from prototype patterns and the input neurons are termed codebook vectors. The third layer is an output layer. The third layer and the weighted links to the output one perform as static linear map converting the class of the code book vectors into the classes presented at the output neurons [29].

There are two reasons why the learning vector quantization algorithm is noticeable: one is heuristic simplicity, the other is direct adaptability to the text classified tasks. Nevertheless, the employment of learning vector quantization algorithm has not been sufficiently developed in text classification. The learning vector quantization method is a classifying means which is the basis of the artificial neural network competitive studying, which permits the concept of a cluster of kinds in the input data space by employing intensive studying, either negative (penalization) or positive (prize).

The architecture of artificial neural network is very simple but efficient. NO hidden layer units belong to the kohonen model which includes one output layer and one input layer. Kohonen network has many input layer units but it has one output layer units.

Each input layer unit contacts each output layer unit by feedforward connections. Nevertheless, there are interrelations between all the output layer units by lateral inhibitory relations (LIR). Therefore, the output layer units can send inhibitory information and deactivates the remainder of the output units by using the lateral relations as an output layer unit is activated.

The weighting vector contacting every output layer unit is termed as a code-book vector in the learning vector quantization algorithm. Each type of the input space is representative of its own suit of code-book vectors. In the light of the specific task we can define the code-book vectors. We employ a code-book vector to represent each class for classification tasks. The class label of each code-book vector indicates the name of each type. We employ as much as possible of the code-book vectors instead of proper nouns for dimension reduction, and thus it can avoid ambiguity. Under the circumstances, the tags of each type are the type names for each code-book vector.

LQV algorithm is a competitive artificial neural network, therefore, for each of the training vector, the output layer unit will contend with each other in order to look for the champion one according to some indicators. Learning vector quantization algorithm employs Euclidean distance to look for the champion of the units. Nothing but champion unit can revise its weight factor by making good use of the studying rule of learning vector quantization.

We conclude the major steps of the standard learning vector quantization algorithm [30]:

Step 1: Initialize the parameters of LVQ, such as code-book vector W_i and studying rate α.

Step 2: Select an input layer vector X randomly.

Step 3: Find the champion unit which is near to the input layer vector :

$$||X-W_c|| = \min_k ||X-W_k|| \tag{9}$$

that is

$$c = arg\min_k ||X-W_k||. \tag{10}$$

Step 4: The weights factor of the champion unit must be Modified:

–On condition that W_c and X pertain to the identical type

$$W_c(t+1) = W_c(t)+\alpha(t)[X(t)-W_c(t)] \tag{11}$$

–On condition that W_c and X pertain to different types

$$W_c(t+1) = W_c(t)-\alpha(t)[X(t)-W_c(t)] \tag{12}$$

Step 5: Perform the reduction of the studying rate α.

Step 6: Perform repetitively the steps from step 2 to step 5 until the termination criterion is met.

The studying rate $\alpha(t)$ $(0<\alpha(t)<1)$ is a drably decreasing equation of time series and it can be employed to regulate the speed of the weight factor which is permitted to alter. We can conclude that $\alpha(t)$ should be at least smaller than 0.25 and it decreases constantly to infinitesimal positive real number which is defined by users.

3 Face recognition method based on the improved LDA and LVQ

In this paper, the learning vector quantization is manipulated by an improved linear discriminant analysis strategy and then applied to solve the face recognition problem.

3.1 An improved linear discriminant analysis strategy

In this paper, the new discrete degree matrix between classes is denoted in the following:

$$M_b^{NEWLDA} = \sum_{i=1}^{L} \frac{\sum_{j=1}^{N} u_{ij}}{N}(c_i-c)(c_i-c)^T \tag{13}$$

Where $c_i = \sum_{j=1}^{N}(u_{ij}/\sum_{k=1}^{N} u_{ik})x_j$ represents the class mean and is different from the traditional c_i in LDA,

$$u_{ij} = \frac{1}{\sum_{k=1}^{L} \frac{||x_j-c_i||^{\frac{2}{m-1}}}{||x_j-c_k||^{\frac{2}{m-1}}}}, u_{ij}\in[0,1] \tag{14}$$

subject to $\sum_{i=1}^{L} u_{ij} = 1$, for $i=1,\cdots,L,\ j=1,\cdots,N$. Thus, based on the Fisher criterionthe optimal evaluation could be described as follows

$$J_{NEWLDA} = tr[(M_w^{LDA})^{-1}M_b^{NEWLDA}] \tag{15}$$

3.2 ILDALVQ Algorithm

There are five steps of the ILDALVQ algorithm as described below:

Step 1: Initialize the parameters, the matrix $U=[u_{ij}]$, $1\le i\le L$, $1\le j\le N$, and the transmission parameter of radial basis function $sp=0.1$, where $u_{ij}\in[0,1]$.

Step 2: Find the optimal value U by the Fuzzy C-means algorithm.

Step 3: Gain the eigenvectors and eigenvalues by the improved LDA algorithm and reduce the dimensions of the training sample and test sample.

Step 4: Train the test sample by the probabilistic neural network and then design the grader.

Step 5: Classify the test samples and output the results.

4 Experimental results and analysis

To facilitate the experiments, we used the Neural Network Toolbox in matlab2012a to program a m-file for implementing the improved LDA and LVQ on a personal computer with a 32-bit windows 7 operating system, a 4GB of RAM, and a 3.10GHz-core(TM) i5-based processor.

The ORL face database includes a sum of 400 face images, there are 40 people and each person has 10 samples. The face images were collected from different subjects with different facial expressions (there are open/closed eyes, glassed/not glassed and smiling/not smiling et al) at various times and different lighting condition. All of the face images were gathered from the same dark background and there are different orientations like upward/downward, forward/backward and oblique/rotary and so on. Each face image which was manually handled contains the picture elements of 32×32 in our study. Some face images of ten emblematical people are demonstrated in Fig. 1.

There are 15 subjects and each subject has 11 gray-scale face images, so the sum of images is 165 in the Yale face database. And the face images in this database were gathered from different subjects with different facial expressions (there are normal/abnormal, sad/happy, awake/sleepy, angry/mild and wink/surprised et al) at various times and different lighting conditions. Each face image which was manually handled contains the picture elements of 32×32 in order to perform the experiment conveniently in our study. Some face images of ten emblematical people are demonstrated in Fig. 2.

There are 126 subjects and the sum of images is over 4000 in the AR database of face images. The face images were obtained from different subjects with different facial expressions (there are open/closed eyes, glassed/not glassed and smiling/not smiling et al) at various times and different lighting conditions (open /occlusions). Each subject has 26 color face images which are originated from two different sections. 1680 face images were selected from the AR face database and were made up of a subset, the subset includes 120 people (there are 60 women and 60 men) and each people has 14 face images. Each face image which was manually handled contains the picture elements of 40×50 in order to perform the experiment conveniently in our study. Some face images of ten emblematical people are demonstrated in Fig. 3.

Table 1 shows the number of samples of each kind from the ORL database of face images and the Yale database of face images. We choose the second, the forth, the sixth and the eighth one of each kind to be the training samples, and the rest face images are applied to be test samples. For the above three examples, ILDALVQ algorithm are compared with the dominant algorithms, PCA [31], LDA [31], NNC [32], CBNNC [33], NNL [34]. The results are shown below.

Table 1: Training samples.

samples	1	2	3	4	5
ORL	10	45	120	210	252
Yale	11	55	165	330	462

Table 2: Average accuracy on Yale.

method	Average accuracy				
	1	2	3	4	5
PCA	NA	42.8	51.1	56.4	60.2
LDA	NA	47.2	64.9	72.9	78.8
ILDALVQ	36.67	62.96	68.33	75.24	81.11

Fig. 1: Images of ten people in ORL

Criteria face image databases including ORL, Yale and AR were selected to perform the simulation

Fig. 2: Images of ten people in Yale

Fig. 3: Images of ten people in AR

Table 3: Average accuracy on ORL.

method	Average accuracy				
	1	2	3	4	5
PCA	NA	66.9	74.1	77.9	80.7
LDA	NA	71.5	74.1	89.4	92.8
ILDALVQ	51.53	78.57	86.2	90.2	92.62

Table 4: Average accuracy on AR.

method	Correct rate (%)			
	2	4	6	8
PCA	69.62	68.45	69.18	66.11
NNC	60.07	57.31	61.12	58.24
CBNNC	59.97	57.31	61.08	59.24
NNL	59.2	57.5	61.75	58.64
ILDALVQ	74.31	70	66.67	85.56

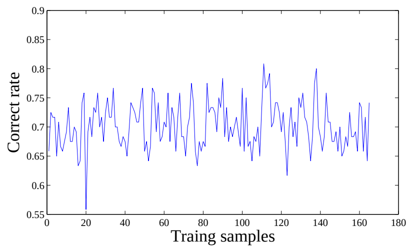

Fig. 4: Correct rate on Yale dataset

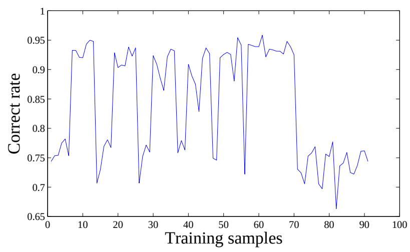

Fig. 5: Correct rate on AR dataset

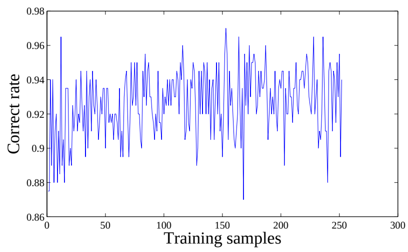

Fig. 6: Correct rate on ORL dataset

experiments. Tables 2, 3 and 4 show the comparison results between this proposed algorithm and other algorithms; Figure 4-6 exhibit the results of the proposed algorithm with Yale database, AR database and ORL database, when each training sample selection and the recognition rate of the reach of the schematic diagram.

By applying the algorithm to Table 2, which considers the Yale face database, it produces discriminant results with an average accuracy, compared with classical classification algorithm PCA and LDA. However it produces a higher accuracy, when each sample size is small, with very good classification effect. For the ORL database in Table 3, despite the number of samples for each 5, this proposed algorithm obtains a slightly low classification effect compared with the traditional LDA algorithm, but in general, the classification effect is better than the PCA and the LDA. For the AR database in Table 4, this proposed algorithm has better classification results in the achievement of results compared with other classical algorithms.

5 Conclusion

In the paper, we utilized a modified LDA and LVQ neural network methods to propose a novel method for solving face recognition problem. Our proposed method can decrease the running time of the whole algorithm, and can improve the performance of the neural network. The experimental results showed that the proposed method can effectively solve the problem of face recognition and have more powerful classifying ability than other algorithms in previous literatures.

Acknowledgement

This paper was partially funded by the Key Program of National Natural Science Foundation of China (Grant No.61133005), and the National Natural Science Foundation of China (Grant Nos.90715029, 61070057, 60603053, 61202109), Key Projects in the National Science & Technology Pillar Program the Cultivation Fund of the Key Scientific and Technical Innovation Project, Ministry of Education of China (Grant No.708066), the Ph.D. Programs Foundation of Ministry of Education of China (Grant No.20100161110019) the Program for New Century Excellent Talents in University (Grant No.NCET-08-0177), the Research Foundation of Education Bureau of Hunan Province, China(Grant No.13C333), the Natural Science Foundation of Zhejiang, China(No.LY12F02019), the public technology application research of Zhejiang, China(No.2011C23130), the Jiaxing Committee of Science and Technology, China(No.2012AY1027).
The authors are grateful to the anonymous referee for a careful checking of the details and for helpful comments that improved this paper.

References

[1] R. Chellappa, C. Wilson and S. Sirohey, Human and machine recognition of faces, a survey Proceedings of the IEEE, **83**, 705-740 (1995).

[2] Lillesand, T. M.; Kiefer, R. W.; Chipman, J. W.; Remote sensing and image interpretation, Remote sensing and image interpretation, (2004).

[3] Lim, Jae S., Two-dimensional signal and image processing, Englewood Cliffs, NJ, Prentice Hall, 710, (1990).

[4] P. Belhumeur, J. Hespanha, and D. Kriegman, Eigenfaces vs. Fisherfaces: recognition using class specific linear projection, IEEE Trans. Pattern Anal. Mach. Intell., **19**, 711-720 (1997).

[5] A. Sharma, and K. Paliwal, A two-stage linear discriminant analysis for face-recognition, Pattern Recognit. Lett., **33**, 1157-1162 (2012).

[6] X. Y. Kong, Q. S. An, H. G. Ma, C. Z. Han, and Q. Zhang, Convergence analysis of deterministic discrete time system of a unified self-stabilizing algorithm for PCA and MCA, Neural Netw, **36**, 64-72 (2012).

[7] X. Chen, J. Yang, D. Zhang and J. Liang, Complete large margin linear discriminant analysis using mathematical programming approach, Pattern Recognit, **46**, 1579-1594 (2013).

[8] M. Loog, R. P. W. Duin and R. Haeb-Umbach, Multiclass linear dimension reduction by weighted pairwise Fisher criteria, IEEE Trans. Pattern Anal. Mach. Intell., **23**, 762-766 (2001).

[9] D. Tao, X. Li, X. Wu and S.J. Maybank, Geometric mean for subspace selection, IEEE Trans. Pattern Anal. Mach. Intell., **31**, 260-274 (2009).

[10] P. Howland, and H. Park, Generalizing discriminant analysis using the generalized singular value decomposition, IEEE Trans. Pattern Anal. Machine Intell., **26**, 995-1006 (2004).

[11] G.F. Lu, J. Zou, and Y. Wang, Incremental complete LDA for face recognition, Pattern Recognit., **45**, 2510-2521 (2012).

[12] J. Yang, and J.Y. Yang, Why can LDA be performed in PCA transformed space?, Pattern Recognit., **36**, 563-566 (2003).

[13] F. Nie, S. Xiang, and C. Zhang, Neighborhood MinMax Projections, in: The 20th International Joint Conference on Artificial Intelligence (IJCAI), 993-998 (2007).

[14] Y. Jia, F. Nie, and C. Zhang, Trace ratio problem revisited, IEEE Trans. Neural Netw., **20**, 729-735 (2009).

[15] S. Pang, S. Ozawa, and N. Kasabov, Incremental linear discriminant analysis for classification of data streams, IEEE Trans. Syst. Man Cybern., Part B: Cybern., **35**, 905-914 (2005).

[16] P. Howland and H. Park, Generalizing discriminant analysis using the generalized singular value decomposition, IEEE Trans. Pattern Anal. Mach. Intell., **26**, 995-1006 (2004).

[17] H. Zhao and P. C. Yuen, Incremental linear discriminant analysis for face recognition, IEEE Trans. Syst., Man, Cybern. B, Cybern., **38**, 201-221 (2008).

[18] T. K. Kim, B. Stenger, J. Kittler, and R. Cipolla, Incremental linear discriminant analysis using sufficient spanning sets and its applications, Int. J. Computer Vision, **91**, 216-232 (2011).

[19] Tamer Shahwan and Raed Said, A Comparison of Bayesian Methods and Artificial Neural Networks for Forecasting Chaotic Financial Time Series, Journal of Statistics Applications & Probability, **1**, 89-100 (2012).

[20] Li-Xin Wang, Adaptive fuzzy systems and control: design and stability analysis, Prentice-Hall, Inc. Upper Saddle River, NJ, USA, (1994).
[21] A. Deng, X. Zhang, J. Tang, L. Zhao and K. Qin, Localization of acoustic emission source based on chaotic neural networks, Appl. Math. Inf. Sci., **6**, 713-719 (2012).
[22] M. R. Girgis, T. M. Mahmoud, H. F. Abd El-Hameed, Z. M. El-Saghier, Routing and Capacity Assignment Problem in Computer Networks Using Genetic Algorithm, Information Science Letters, **2**, 13-25 (2013).
[23] Bo Zhang, Zhicai Juan, Modeling User Equilibrium and the Day-to-day Traffic Evolution based on Cumulative Prospect Theory, Information Science Letters, **2**, 9-12 (2013).
[24] T. Kohonen, Self-organizing Maps. Springer, Berlin, (1995).
[25] B. Hammer, M. Strickert, and T. Villmann, On the generalization capability of GRLVQ networks, Neural Process. Lett., **21**, 109-120 (2005).
[26] B. Hammer, and T. Villmann, Generalized relevance learning vector quantization, Neural Netw., **15**, 1059-1068 (2002).
[27] M. Biehl, A. Ghosh, and B. Hammer, Learning vector quantization: The dynamics of winner-takes-all algorithms, Neurocomputing, **69**, 660-670 (2006).
[28] G. F. Lu and Y. Wang, Feature extraction using a fast null space based linear discriminant analysis algorithm, Inf. Sci., **193**, 72-80 (2012).
[29] F. Dieterlea, S. Müller-Hagedornb, H.M. Liebichb and G. Gauglitza, Urinary nucleosides as potential tumor markers evaluated by learning vector quantization, Artif Intell Med, **28**, 265-279 (2003)
[30] M. T. Martín-Valdivia, L.A. Ureña-López, and M. García-Vega, The learning vector quantization algorithm applied to automatic text classification tasks, Neural Netw., **20**, 748-756 (2007).
[31] I. Jlliffe, Principal Component Analysis. New York: Springer-Verlag, (1986).
[32] D. Cai, X. He, Y. Hu, J. Han, and T. Huang, Learning a Spatially Smooth Subspace for Face Recognition, CVPR'07, (2007).
[33] Q. B Gao, and Z. Z. Wang, Center-based nearest neighbor classifiers for pattern classification, Pattern Recognit. **40**, 346-349 (2007).
[34] W. Zheng, L. Zhao, and C. Zou, Locally nearest neighbor classifiers for pattern classification, Pattern Recognit., **37**, 1307-1309 (2004).

Aijia Ouyang received the M.E. degree in the department of computer science and technology from Guangxi University for Nationalities, Nanning, China, in 2010. He is a Lecturer in the College of Computer at Jiaxing University, China. He is currently a Ph.D. candidate in the College of Information Science and Engineering at Hunan University, Changsha, China. His research interests include parallel computing, cloud computing and artificial intelligence. He has published research articles in international conference and journals of parallel computing and swarm intelligence algorithms.

Kenli Li received the Ph.D. in computer science from Huazhong University of Science and Technology, China,in 2003, and the M.Sc. in mathematics from Central South University, China, in 2000. He was a visiting scholar at University of Illinois at Champaign and Urbana from 2004 to 2005. Now He is a professor of Computer science and Technology at Hunan University, associate director of National Supercomputing Center in Changsha, a senior member of CCF. His major research includes parallel computing, Grid and Cloud computing, and DNA computer. He has published more than 70 papers in international conferences and journals, such as IEEETC, TPDS, JPDC, PC, ICPP, and CCGrid.

Xu Zhou received the M.E. degree in the Dept. of Computer at Hunan University, China 2009. She is currently a Ph.D. candidate in the Dept. of Computer at Hunan University, China. Her research interests include DNA computing. She has published research articles in international conference and journals of DNA computing.

Yuming Xu received the M.E. degree in computer from Hunan University, China, in 2009. He is currently a Ph.D. candidate in the Dept. of Computer at Hunan University, China. His research interests include parallel computing. He has published research articles in international conference and journals of parallel computing.

Guangxue Yue is a Professor in the Dept. of Computer at Jiaxing University, China. He received the Ph.D. degree in the Dept. of Computer from Hunan University, China, 2012. His research interests include computer network and distributed system. He has published more than 70 international journals/conference papers on computer network and information security.

Lizhi Tan is an Assistant Professor in the Dept. of Electronics Engineering at Zhuzhou Professional Technology College, China. He received the M.E. degree in Circuits and systems from the Dept. of Electron Engineering, Hunan University, China, in 2010. His research interests include neural network and pattern recognition. He has published research articles in international conference and journals of electronics engineering and neural network.