Recognition algorithm for plant leaves based on adaptive supervised locally linear embedding

Yan Qing[1], Liang Dong[1], Zhang Dongyan[1,2*], Wang Xiu[2]

(1. *Key Laboratory of Intelligent Computing & Signal Processing, Ministry of Education, Anhui University, Hefei 230039, China*; 2. *Beijing Research Center for Information Technology in Agriculture, Beijing 100097, China*)

Abstract: Locally linear embedding (LLE) algorithm has a distinct deficiency in practical application. It requires users to select the neighborhood parameter, *k*, which denotes the number of nearest neighbors. A new adaptive method is presented based on supervised LLE in this article. A similarity measure is formed by utilizing the Fisher projection distance, and then it is used as a threshold to select *k*. Different samples will produce different *k* adaptively according to the density of the data distribution. The method is applied to classify plant leaves. The experimental results show that the average classification rate of this new method is up to 92.4%, which is much better than the results from the traditional LLE and supervised LLE.

Keywords: supervised locally linear embedding, manifold learning, Fisher projection, adaptive neighbors, leaf recognition, Precision Agriculture

DOI: 10.3965/j.ijabe.20130603.007

Citation: Yan Q, Liang D, Zhang D Y, Wang X. Recognition algorithm for plant leaves based on adaptive neighborhood optimization supervised locally linear embedding. Int J Agric & Biol Eng, 2013; 6(3): 52－57.

1 Introduction

Leaf recognition is one of the effective ways to recognize plant species. The pattern recognition method can improve the efficiency of leaf recognition to a large extent. Scholars have obtained some achievements in this field. The methods generally extract biological characteristics of leaves, such as color, shape, or texture, and then use a classifier to recognize them[1-7]. However, the extracted features are often of high dimensions, so the dimensionality reduction is necessary before carrying out the classification. The traditional dimensionality reduction methods, such as principal component analysis (PCA)[8] and independent component analysis (ICA)[9], are linear. As leaves features can be easily influenced by natural conditions, such as lighting and distortion, the data distribution is often actually nonlinear. Therefore linear dimensionality reduction methods cannot maintain the internal topology of the data, which would influence the recognition result directly. At present, manifold learning is the research focusing on nonlinear dimension reduction. Locally linear embedding (LLE) is one of the manifold learning algorithms, which has been applied widely in image processing[10], because it can find the globally optimal solution and its computation complexity is low. However, the traditional LLE method still has some unsolved problems in realization and the selection of parameter, for example, *k*, the number of the nearest neighbors. The LLE method is sensitive to the value of *k*, and it is not reasonable to use a uniform *k* to every sample because usually the distribution of data is not uniform.

Owing to the problem mentioned above, many experts

Received date: 2013-04-17 **Accepted date:** 2013-07-01

Biographies: Yan Qing, PhD student, Lecturer; Research interests: digital image processing and pattern recognition; Email: rubby_yan5996@sina.com. **Liang Dong,** PhD, Professor; Research interests: computer vision and digital signal processing; Email: dliang@ahu.edu.cn. **Wang Xiu,** PhD, Professor; Research interests: intelligent machinery of precision agriculture and automatic control; Email: wangx@nercita.org.cn.

*** Corresponding author: Zhang Dongyan,** PhD; Research interests: computer image processing, hyperspectral remote sensing technology and kinds of sensors applying in agriculture and environment; Email: hello-lion@hotmail.com.

have also proposed some adaptive methods based on unsupervised LLE, while these methods often bring in some new parameters[11-17]. When LLE is directly applied to solve the classification problem, the recognition rate will be influenced because it cannot adequately utilize the category information of the learning samples. Improved supervised LLE has achieved some in leaf recognition[18,19]. However supervised LLE[19] still presents the same problem mentioned above. As it is well known, the selection of neighborhood parameter, *k*, has a large influence on the recognition rate. The study finds that if the parameter *k* is oversized, it will probably include some samples coming from different category in certain sample's neighborhood, which will provide inaccurate information and destroy the manifold structure. The continuous manifold structure will be divided into many sub-manifolds if *k* is undersized. In order to solve this problem, an algorithm based on supervised LLE is proposed which can adaptively determine the value of *k*. The algorithm constructs a similarity index using the Fisher projection distance, which can select the value of *k* according to actual distribution of samples. When applying the algorithm to leaf recognition and comparing it with supervised LLE, the experimental results prove that the proposed method can improve the classification accuracy with a reasonable selection of neighborhood parameter *k*.

2 Supervised LLE algorithm based on the Fisher criterion (FS-LLE)

Traditional LLE is an unsupervised algorithm, and it constructs the neighborhood structure on the basis of the Euclidean distance. The Euclidean distance simply considers that the dimension of the data is irrelevant with each other, which is not correct for the image data. Hence, if we adopt the Euclidean distance to construct the neighborhood of each sample, the result is perhaps different from human's perceptions. Considering the disadvantage of the unsupervised LLE, a supervised LLE is constructed as follows adopting the Fisher criterion which can explore category information[19].

1) It is assumed that the dataset after sampling is $X\{x_1, x_2, ...x_n \in R^D\}$, and each sample point x_i is projected by the Fisher criterion to the Fisher space. Here, *n* is the number of samples and *D* is the original dimension of each sample. According to the distances between the projection point $\tilde{x}_i$, the nearest neighbors of each original sample, *k*, is determined.

2) Minimize (1) to calculate the reconstruction weight w_{ij}, then reconstruct sample point x_i by its nearest neighbors, *k*:

$$\varepsilon(W) = \sum_i \left| x_i - \sum_{j=1}^{k} w_{ij} x_j \right|^2 \tag{1}$$

where, $\varepsilon(W)$ is the reconstuction cost function, and weight w_{ij} meets the constraint condition $\sum_j w_{ij} = 1$, and $w_{ij} = 0$ if x_i is not a neighbor of x_j.

3) A low-dimensional vector *Y* is constructed by keeping w_{ij} unchanged and then the following error function is minimized:

$$\phi(Y) = \sum_{i=1}^{n} \left\| y_i - \sum_{X_j \in \Omega(X_i)} w_{ij} y_j \right\|^2 = \left\| Y(I - W^T) \right\|^2 = tr(YNY^T) \tag{2}$$

where, $y_i \in R^d (d << D)$, $\sum_i y_i = 0$, and $\frac{1}{n}\sum_i y_i y_i^T = I$.

So the coordinates of the low dimensional manifold are the smallest feature vector from 2 to *d*+1 of the matrix $N = (I - W)^T(I - W)$, *d* is the dimensions of the low dimensional manifold *Y*.

The supervised LLE can fully use the supervised information of training samples, so it improves the recognition accuracy effectively. However, the value of *k* is obtained by doing the experiment many times and the selection lacks guidance. Even more important is that the selection of *k* will influence the recognition accuracy directly.

3 Supervised LLE algorithm with adaptive neighborhood optimization

3.1 Adaptive neighborhood locally linear embedding algorithm (ANLLE)

Zhang and Huang[17] proposed the ANLLE algorithm, which can select the number of neighbors adaptively. The algorithm defines the similarity of any two samples x_i and x_j as:

$$S_{ij} = 1/\| x_i - x_j \| \quad (3)$$

The average similarity between x_i and the other samples is:

$$M_i = \frac{1}{n-1}\sum_j \frac{1}{\| x_i - x_j \|}, j \le N \text{ and } j \ne i \quad (4)$$

The M_i is selected as the threshold to determine the value of k, if the similarity of two random x_i and x_j is greater than M_i, the x_i is the neighbor of x_j. Zhang and Huang[17] proved that this method not only can select the value of k automatically, but also can improve the classification accuracy effectively. Nevertheless, this similarity measure is still based on the Euclidean distance, so it does not utilize the supervised information effectively, and the classification accuracy is limited when it is used for classification directly.

3.2 Supervised LLE algorithm with adaptive neighborhood optimization based on Fisher projection (FS-ANLLE)

As described above, the Fisher transform determines the best projection direction for optimum classification, so the projection distance in the determined direction could reflect the class difference of the training sample and improve the clustering ability of samples, furthermore to improve the recognition accuracy.

Yan et al.[19] proposed a supervised LLE algorithm based on Fisher transform which is called FS-LLE, but this algorithm still depends on manual selection of k. This study makes an improvement based on the FS-LLE and ANLLE. The projection distance is used to modify the similarity measure in Equation (4) and select the adaptive neighborhood parameter depending on the similarity. The detail and procedure of the algorithm are as follows:

1) Take Fisher projection for each sample, then utilize the projection distance to make required transforms in Equations (3) and (4), which results in the following, new similarity and average similarity.

$$S^{\mathrm{f}}{}_{ij} = 1/ FisherD(x_i, x_j) \quad (5)$$

$$M^{\mathrm{f}}{}_i = \frac{1}{n-1}\sum_j \frac{1}{FisherD(x_i, x_j)}, j \le n \text{ and } j \ne i \quad (6)$$

where, $FisherD(x_i, x_j)$ is the projection distance among samples.

2) The dichotomy is used to determine the number of neighbors of each sample x_i. The flow chart is shown in Figure 1.

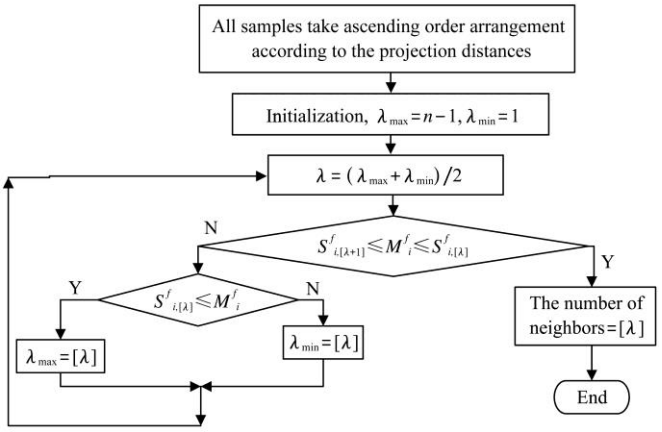

Figure 1 Dichotomy flow chart

3) This step is the same as step (2) of FS-LLE. It depends on the neighbors of each sample point to calculate the reconstruction weights, and reconstruct the original samples by their neighbors according to Equation (1). The difference is the number of all samples' neighbors equal in FS-LLE, while each sample selects the number of neighbors automatically according to the result of the two steps above in this algorithm. Hence, the neighborhood structure is established by the actual samples distribution.

4) The low dimensional embedding coordinates are obtained in the same way as the step (3) of FS-LLE.

5) The same method is used to calculate the similarity matrix and average similarity of each test sample. Then the k_{test}, which is the number of neighbors of each test sample, is determined. The test sample is reconstructed by its neighbors, and then the reconstruction weights are obtained, and finally the unchanged weights remain. The test sample's low dimensional embedding coordinate is approximated by the low dimensional coordinate of training samples.

4 Simulation experiment and analysis

The database (http://www.intelengine.cn/data) used in the experiment is provided by Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei, China. The experiment uses dogbane oleander (*Nerium indicum Mill.*) leaves as the positive training samples, and other

leaves' as the negative training samples. Some positive and negative leaf images are shown in Figure 2.

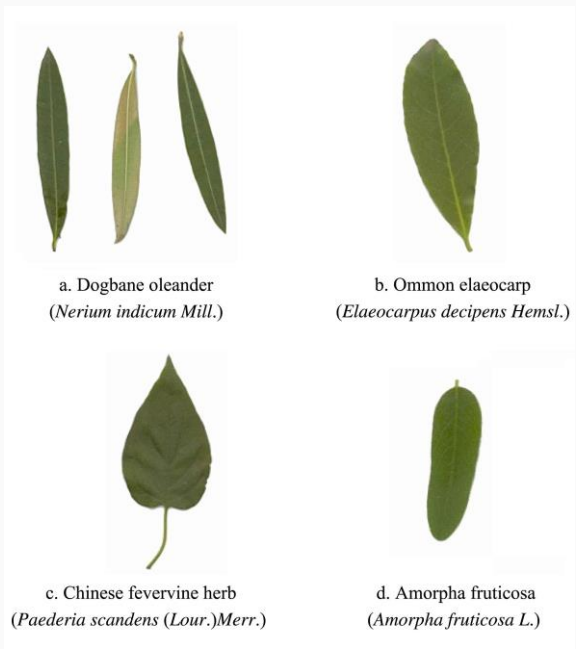

Figure 2 Positive and negative leaf images

4.1 Image preprocessing

Each leaf image is colored and has different sizes in this database, so it is necessary to transform the images into grey images and reduce them to 64 × 64 pixels by wavelet transformation before the experiment. Each sample can be described by a square matrix, and then all square matrixes are reshaped to a 4096-dimensional row vector.

4.2 Cluster analysis

The clustering effect can reflect the effectiveness of the dimension reduction method for classification. The first experiment was designed to test the clustering affected by FS-ANLLE. Thirty positive samples and 60 negative samples were chosen as dataset **I**, then LLE, FS-LLE, and FS-ANLLE were used to reduce the dimensions. The three dimensional space distributions are shown in Figure 3.

The clustering figure shows that LLE has an unsatisfactory clustering effect in low dimensional space, because LLE is an unsupervised algorithm and does not use the data category information in dimension reduction, so LLE would restrict the classification accuracy. FS-LLE uses the supervised information fully and its clustering ability has improved significantly. The FS-ANLLE proposed in this paper is based on FS-LLE, and it not only avoids the influence brought by unreasonable selection of the neighbourhood parameter, but also improves the clustering effect of low dimensions further.

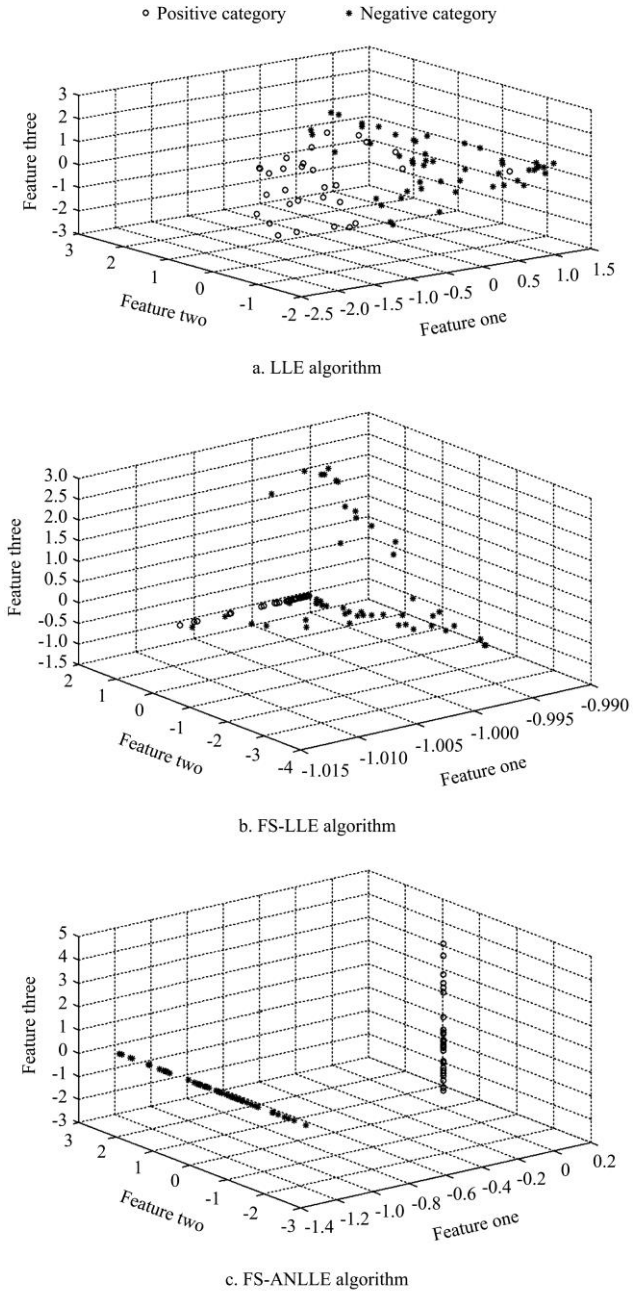

Figure 3 Comparison of the clustering effects of different algorithms

4.3 Classification experiment

In classification experiment, 38 dogbane oleander leaves and 20 other kinds of leaves were selected as test samples. In latter experiment, the LLE, FS-LLE and FS-ANLLE were used as dimension reduction method for the comparative purpose, then the classification was conducted by the nearest neighbor classifier.

In fact there is another important undetermined parameter, the dimension *d*, which is low dimensional manifold. The second experiment was carried on, in which the dataset I was still selected. The purpose of this experiment was to find out the optimal parameters, *k* and *d*, in the method of LLE. Thus we could compare FS-ANLLE with the other two algorithms with optimal parameters, so this comparison is more meaningful. The experimental results are shown in Table 1.

Table 1 The classification rates of LLE when the values of *k* and *d* are changed (%)

	k = 6	*k* = 8	*k* = 10	*k* = 12	*k* = 14
d = 5	86.2	84.5	86.2	86.2	86.2
d = 10	86.2	86.2	86.2	89.7	89.7
d = 15	84.5	84.5	84.5	86.2	89.7
d = 20	84.5	82.8	82.8	84.5	84.5
d = 25	82.8	82.8	82.8	84.5	82.8

From Table 1, the selection of *k* and *d* can influence the accuracy of recognition of leaf images. When *k* was 12, the classification rates were ideal along with the change of *d*, so in the third experiment the LLE and FS-LLE were carried on under the condition of *k* of 12. This experiment was iterated 20 groups, and in each group we chose 30 positive samples and 60 negative samples randomly as training samples from database. Then 38 positive samples and 20 negative samples were chosen as the test samples. The three algorithms were used to reduce the sample's dimensions, and then the nearest neighbor classifier was used to classify the samples. Table 2 shows the best classification rate of the different algorithms in the 20 group experiment with different *d* values.

Table 2 Comparison of the classification results of the three methods (%)

	LLE	FS-LLE	FS-ANLLE
d = 5	86.2	91.34	93.1
d = 10	89.7	93.1	91.4
d = 15	86.2	93.1	94.8
d = 20	84.9	93.1	93.1
d = 25	84.9	89.7	89.7
Average classification rate	86.2	92.1	92.4

From Table 2, it is clear that the classification of FS-LLE and FS-ANLLE are better than that of LLE because they use supervised information reasonably. Comparing FS-ANLLE with FS-LLE, the average classification rate has been improved further; only when *d* is 10, the classification rate of FS-ANLLE (91.4%) is a little lower than that of FS-LLE (93.1%). The advantages of FS-ANLLE are that it avoids subjective selection of the parameter *k*, improves the algorithm's implementation efficiency, and, at the same time, improves the classification accuracy to an extent because the selection of *k* is more consistent to the actual samples distribution.

5 Conclusions

The method proposed by this study uses the projection distance to construct the sample's similarity index and selects number of neighbors. The method not only realizes the adaptive supervised LLE algorithm, but also obtains a better classification result. A future research should focus on how to adaptively select the parameter other than *k*.

Acknowledgements

This study was financially supported by the National Natural Science Foundation of China (61172127), the Research Fund for the Doctoral Program of Higher Education(KJQN1114), Anhui Provincial Natural Science Foundation (1308085QC58), the 211 Project Youth Scientific Research Fund of Anhui University, and Provincial Natural Science Foundation of Anhui Universities (KJ2013A026)

[References]

[1] Du J X, Huang D S, Wang X F, Gu X. Shape recognition based on radial basis probabilistic neural network and application to plant species identification. Lecture Notes in Computer Science, 2005; 3497: 281-285.

[2] Gu X, Du J X. Leaf recognition based on the skeleton segmentation. Lecture Notes in Computer Science, 2005; 3644: 253-262.

[3] Du J X. Study of plant leaf recognition techniques by machine. University of Science and Technology of China, PhD thesis, 2005, China. (In Chinese with English abstract)

[4] Li Y F, Zhu Q S, Cao Y K, Wang C L. A leaf vein extraction method based on snakes technique. Proceedings of IEEE International Conference on Neural Networks and Brain, 2005: 885-888.

[5] Neto J C, Meyer G E, Jones D D, Samal A K. Plant species identification using Elliptic Fourier leaf shape analysis. Computers and Electronics in Agriculture, 2006; 50(2): 121-134.

[6] Bruno O M, Plotze R O, Falvo M, de Castro M. Fractal dimension applied to plant identification. Information Science, 2008; 178(12): 2722-2733.

[7] Wang X F, Huang D S, Du J X, Zhang G J. Feature extraction and recognition for leaf images. Computer Engineering and Applications, 2006; 2006(3): 190-193.

[8] Chen F B, Yang J Y. Modular PCA and its application in human face recognition. Computer Engineering and Design, 2007; 28(8): 1889-1892.

[9] Wang H M, Ou Z Y. Face recognition based on features by PCA/ICA and classification with SVM. Journal of Computer Aided Design & Computer Graphics, 2003; 15(4): 416-420. (In Chinese with English abstract)

[10] Roweis S T, Saul L K. Nonlinear dimensionality reduction by locally linear embedding. Science, 2000; 290(5500): 2323-2326.

[11] Li B, Yang D, Lei M, Ge Y X. Adaptive locally linear embedding based on affinity propagation. Journal of Optoelectronics Laser, 2010; 21(5): 772-778.

[12] Wen G H, Jiang L J, Wen J. Locally linear embedding based on optimization of neighborhood. Journal of System Simulation, 2007; 19(13): 3119-3122. (In Chinese with English abstract)

[13] Wen G H, Jiang L J, Wen J. Dynamically determining neighborhood parameter for locally linear embedding. Journal of Software, 2008; 19(7): 1666-1673. (In Chinese with English abstract)

[14] Yu J, Qin R X, Deng N Y. Locally linear embedding algorithm based on adaptive nearest neighbor. Control Engineering of China, 2006; 13(5): 469-470. (In Chinese with English abstract)

[15] Hui K H, Xiao B H, Wang C H. Self-regulation of neighborhood parameter for locally linear embedding. Pattern Recognition and Artificial Intelligence, 2010; 23(6): 842-846. (In Chinese with English abstract)

[16] Zhang Y L, Zhuang J, Wang N, Wang S A. Fusion of adaptive local linear embedding and spectral clustering algorithm with application to fault diagnosis. Journal of Xi′an Jiaotong University, 2010; 44(1): 77-82. (In Chinese with English abstract)

[17] Zhang X F, Huang S B. Adaptive neighborhoods based locally linear embedding algorithm. Journal of Harbin Engineering University, 2012; 33(4): 489-495. (In Chinese with English abstract)

[18] Zhang S W, Huang D S. A robust supervised manifold learning algorithm and its application to plant leaf classification. Pattern Recognition and Artificial Intelligence, 2010; 23(6):836-841. (In Chinese with English abstract)

[19] Yan Q, Liang D, Zhang J J. Recognition method of plant leaves based on Fisher projection-supervised LLE algorithm. Transactions of the Chinese Society for Agricultural Machinery, 2012; 43(9): 179-183. (In Chinese with English abstract)