An Elliptical Basis Function Network for Classification of Remote-Sensing Images

Jian-Cheng Luo Author1
Qiu-Xiao Chen Author2
Jiang Zheng Author3
LREIS, Institute of Geographical Science and Natural Resources Research,
Chinese Academy of Sciences, Beijing, China, 100101
E-mail: luojc@lreis.ac.cn

Yee Leung Author4
Department of Geography and Resource Management,
The Chinese University of Hong Kong, Hong Kong

Jiang-Hong Ma Author5
Faculty of Sciences, Xi' an Jiaotong University, Xi' an, China, 710049

***Abstract*-An elliptical basis function (EBF) network is proposed in this study for the classification of remotely sensed images. Though similar in structure, the EBF network differs from the well-known radial basis function (RBF) network by incorporating full covariance matrices and uses the expectation-maximization (EM) algorithm to estimate the basis functions. Since remotely sensed data often take on mixture-density distributions in the feature space, the proposed network not only possesses the advantage of the RBF mechanism but also utilizes the EM algorithm to compute the maximum likelihood estimates of the mean vectors and covariance matrices of a Gaussian mixture distribution in the training phase. Experimental results show that the EM-based EBF network is faster in training, more accurate, and simpler in structure.**

***Keywords*-artificial neural networks, classification, elliptical basis functions, EM algorithm, mixture densities, radial basis functions, remotely sensed image**

1. INTRODUCTION

Artificial neural networks (ANN) have been successfully applied in information extraction and classification of remotely sensed data. Among various kinds of ANN architectures, radial basis function (RBF) networks appear to be a more effective model for nonlinear function approximation and data classification in general, and remote-sensing classification in particular [3][5][7].

It is well known that the classification error made by the RBF networks depends strongly on the selection of the centers and widths of the kernel functions constituting the hidden layer [1][4]. Conventional clustering algorithms, such as K-means and K-nearest neighbors, are not very effective for the high computation cost on mixture density estimation caused by symmetric property of basis functions.

In recent years, the application of the EM algorithm in the estimation of parameters of probability density functions has been paid great attention. The EM algorithm is an iterative method, which can be used to numerically approximate the maximum likelihood (ML) estimates of the parameters in a mixture model [2][6], for example, the mean vectors and covariance matrices of a Gaussian mixture distribution. EM is especially fit for the statistical analysis of remote-sensing data, characterized by a mixture with normally distributed components, for its data-based statistical properties rather then heuristic methods.

To take advantages of the efficient structure of the RBF networks and the effective iterative mechanism of the EM algorithms, a new model, similar to the original probabilistic neural networks, is proposed in this paper, called EBF networks for the classification of remotely sensed images. Incorporating full covariance matrices with RBF and employing the EM algorithm, on one hand, EBF possess good ability on approximation for the hyper-ellipsoidal character of its hidden layer, in addition, better result of parameters estimation can be obtained for the data-based character of EM algorithm. Finally, a least square minimization is introduced to determine the output weights.

In this paper, the universal EM algorithm for the estimation of parameters of the mixture distributions is first proposed, and then, structure of EBF network is described as well as the procedure of estimation of the Gaussian parameters with EM algorithm. In addition, detail of EM-based EBF network for the classification of remotely sensed images and a case study are also provided.

2. ESTIMATES OF PARAMETERS OF MIXTURE DENSITY DISTRIBUTIONS IN REMOTELY SENSED DATA USING EM

One of the characters of remotely sensed data is mixture distribution on feature space. In most cases, it is very difficult to compute the MLE(maximum likelihood estimates) by conventional methods directly for the complexity of the mixture density functions.

Similar to pattern recognition in general, one of the key problems of information extraction and classification in remote sensing is the detection and description of serial procedures, phenomena and objects hidden in data sets. Such information is often regulated by finite mixture feature sets with different distribution properties. Toward those conventional, EM is an ideal alternative, which numerically approximate the MLEs of the parameters in an iterative manner.

In most cases, the EM algorithm can always estimate the optimized estimate of the MLEs with iterative procedure. And based on the application of EM for estimating the parameters of the basis functions in EBF hidden layer, bridge related mixture density distributions to supervised labeling categories for the unknown patterns can be estabilished.

3. EBF NETWORKS CLASSIFICATION MODEL

3.1 EBF Network and its EM algorithm

Compared to the original RBF network, the most important improvement of EBF is the introduction of full covariance matrices to basis functions for the representation of complex mixture densities distribution, which is responsible for establishing the connection between input layer and hidden layer. In addition, the connective structure of the linear perceptron is again adopted in network between the hidden layer and the output layer.

Based on the convention methods for EBF networks, EM algorithm is introduced to training the network for better result, where the main function of EM algorithm is to determine the status of the network hidden units, including the number of cluster centers and the means and covariance matrices of the centers. Included steps can be outlined as follow:

(1)Selection of sample data sets.
(2)Initialization. Determine the initial cluster number, the mean vector and covariance matrix of each class.
(3)Estimation of the Maximum Likelihood parameters.
(4)Determination of the optimal number of cluster centers of the EBF network.

In a word, the EBF network can determine the degree of closeness between the input vector and the hidden center with the elliptical response from the corresponding basis function, which makes it more suitable for classification of remotely sensed data compared with conventional RBF network.

3.2 EBF-based Remotely Sensed Image Classification

The basic mechanism of EBF classification network is composed of two parts. First, using the EM algorithm to decompose the mixture density distributions of remotely sensed data in feature space into hidden centers represented with elliptical spreads of probability distribution functions. And then, using the linear perceptron to establish the approximate relationship between cluster centers layer and categorical classes.

On the implementation level, the network consists of three major components, which are the EM algorithm module, the EBF network training module and the classification module respectively.

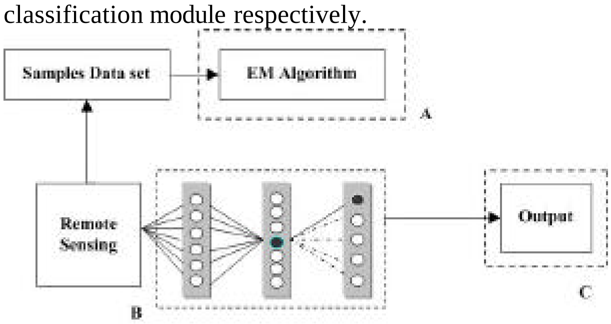

Fig. 1. Architecture of the EM-based EBF Classification Network.

4. Applicaton

4.1 The Study Area and experimental data

To evaluate the performance of the proposed EBF network, a case application of land covers classification with remotely sensed data was made in this study. The experiment was conducted using *SPOT-HRV* data, acquired on *Feb* 3, 1999 covering the Hong Kong island, with a size of 600 rows by 800 columns, included all three spectral bands.

Through visual interpretation of the corresponding remote-sensing data, nine main classes of land covers are identified as the reference for classification, included: Sea Water, Inland Water, Urban Area, Concrete Land, Baren Land, Beach, Grass Land, Hilly Woodland, Grass Land after fire.

Typical sample data sets of each land cover are respectively selected for supervised classification. A total of 3500 supervised samples were selected for the training phase with 2600 training sample data and 900 test sample data.

4.2 Experimental Results and Comparison

With 2600 training samples, mixture densities of the three-dimensional feature space are decomposed into 62 clusters by the EM algorithm, and the ML parameters of each cluster are estimated at the same time. Therefore, the size of the hidden layer of the EBF network is of 62 nodes. In the linear training phase from the hidden layer to the output layer, the training rate h is kept to 0.02, small enough to avoid vibration. The test error matrix is obtained by using the left 900 samples in the trained EBF network. The time for the training phase of the EBF network is about 120 seconds, and the overall test accuracy is up to 76.00%.

As a comparison, the conventional RBF Network (mixture density model, but hyper-spherical) classification models are also trained and tested with the same sample data sets. The c-means clustering algorithm is used to determine the size of the hidden layer of the RBF network resulting in 64 nodes (keeping a scale similar to the EBF network). The test error matrices obtained from the RBF model is listed in Table 1. The overall test accuracy of the RBF network is 70.33%. The training phase of the RBF network is about 50 seconds.

Table 1. Error Matrix of classification by the RBF Network

	C1	C2	C3	C4	C5	C6	C7	C8	C9	Total
C1	91	22	1	0	0	0	0	0	0	114
C2	6	55	9	1	0	1	0	0	0	72
C3	2	22	76	4	0	2	0	1	7	114
C4	0	0	6	67	4	5	0	1	14	97
C5	0	0	0	9	18	3	0	0	0	30
C6	1	0	0	7	78	82	2	0	0	170
C7	0	0	0	4	0	5	81	0	0	90
C8	0	0	2	3	0	2	12	94	9	122
C9	0	1	6	5	0	0	5	4	70	91
Total	100	100	100	100	100	100	100	100	100	900

(Training Time =50Sec, Accuracy = 70.33%, Kappa = 0.666)

Table 2. Error Matrix of classification by the EBF network

	C1	C2	C3	C4	C5	C6	C7	C8	C9	Total
C1	97	18	2	0	0	0	0	0	0	117
C2	2	62	7	0	0	1	0	0	0	72
C3	1	20	76	4	0	0	3	1	6	111
C4	0	0	10	71	6	7	0	0	12	106
C5	0	0	0	7	49	8	0	0	0	64
C6	0	0	0	7	44	71	2	0	0	124
C7	0	0	0	4	1	9	88	1	0	103
C8	0	0	1	2	0	3	5	91	1	103
C9	0	0	4	5	0	1	2	7	81	100
Total	100	100	100	100	100	100	100	100	100	900

(Training Time =120Sec, Accuracy = 76.11%, Kappa = 0.731)

Comparing the two classifiers, following conclusions can be summarized:

(1) The EBF network keeps the major advantages of the RBF network. Though data are freely distributed, the EBF network is more capable of separating the categories of mixture distribution in the feature space.

(2) With similar the size of the hidden layer, the EBF network can attain higher level of accuracy in classification than the RBF network.

So, the EBF network yields the most accurate classification result in both the training and testing phases.

In addition, classification ability of the EBF network in different hidden layer size is also tested with mixture data set. Here, different sizes of the hidden layer, including 20, 30, 40, 50, 60, 80 and 100 centers, are selected for evaluation. Table 4 shows an increasing trend of overall classification accuracy with increasing size of the hidden layer, however, the computation time also increases as well. In addition, this trend would terminate after a certain size.

Anyway, for the incorporation of the full covariance matrices into the basis functions of the EBF network, complex mixture density distributions can be represented in a relative simple way. Therefore, in most cases, the EBF network outperforms the RBF network under similar situations.

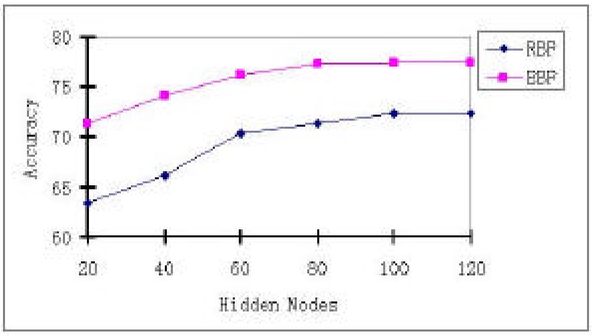

Fig. 2. Comparison of average accuracy between the EBF and the RBF networks.

5. Conclusion

Extending on the structure of the RBF network, an elliptical basis function (EBF) network with full covariance matrices incorporated into the radial basis functions (RBF) whose parameters are estimated by the EM algorithm has been proposed for the classification of remotely sensed images.

To make the proposed EBF network more effective, further research should be carried out to take advantage of the EM algorithm by integrating prior knowledge via Bayes theory, as well as integrating with robustness statistic theory.

REFERENCES

[1]. Bruzzone, L., and D.F. Prieto, 1999b, "A technique for the selection of kernel-function parameters in RBF neural networks for classification of remote-sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, 37(2):551-559.

[2]. Dempster, A.P., N.M. Laird, and D.B. Rubin, "Maximum likelihood estimation from incomplete data via EM algorithm", *J. R. Statist. Soc.*, B39:1-38, 1977.

[3]. Fu, L., 1994, Neural Networks in Computer Intelligence, McGRAW-HILL International Editions.

[4]. Gomm, J.B., and D. Yu, "Selecting Radial Basis Function Network Centers with Recursive Orthogonal Least Squares Training," *IEEE Transactions on Neural networks*, 11(2):306-314, 2000.

[5]. Haykin, S., 1999, Neural Networks: A Comprehensive Foundation, Prentice-Hall.

[6]. McLachlan, G.J., and K.E. Basford, 1988, Mixture Models: Inference and Applications to Clustering, Marcel Dekker, New York.

[7]. Scholkopf, B., K. Sung, C.J.C. Burges, F. Girosi, P. Niyogi, T. Poggio, and V. Vapnik, 1997, Comparing support vector machines with Gaussian kernels to radial basis function classifiers, IEEE Transactions on Signal Processing, 45(11):2758-2765.