Chin. J. Astron. Astrophys. Vol. 3 (2003), No. 3, 277–286
( http://www.chjaa.org or http://chjaa.bao.ac.cn )

Chinese Journal of
Astronomy and
Astrophysics

# Automated Separation of Stars and Normal Galaxies Based on Statistical Mixture Modeling with RBF Neural Networks *

Dong-Mei Qin [1], Ping Guo [2], Zhan-Yi Hu [1] and Yong-Heng Zhao [3]

[1] National Laboratory of Pattern Recognition Laboratory, Institute of Automation, Chinese Academy of Sciences, Beijing 100080; dmqin@nlpr.ia.ac.cn
[2] Department of Computer Sciences, Beijing Normal University, Beijing 100875
[3] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012

Received 2003 January 22; accepted 2003 March 28

**Abstract** For LAMOST, the largest sky survey program in China, the solution of the problem of automatic discrimination of stars from galaxies by spectra has shown that the results of the PSF test can be significantly refined. However, the problem is made worse when the redshifts of galaxies are not available. We present a new automatic method of star/(normal) galaxy separation, which is based on Statistical Mixture Modeling with Radial Basis Function Neural Networks (SMM-RBFNN). This work is a continuation of our previous one, where active and non-active celestial objects were successfully segregated. By combining the method in this paper and the previous one, stars can now be effectively separated from galaxies and AGNs by their spectra—a major goal of LAMOST, and an indispensable step in any automatic spectrum classification system. In our work, the training set includes standard stellar spectra from Jacoby's spectrum library and simulated galaxy spectra of E0, S0, Sa, Sb types with redshift ranging from 0 to 1.2, and the test set of stellar spectra from Pickles' atlas and SDSS spectra of normal galaxies with SNR of 13. Experiments show that our SMM-RBFNN is more efficient in both the training and testing stages than the BPNN (back propagation neural networks), and more importantly, it can achieve a good classification accuracy of 99.22% and 96.52%, respectively for stars and normal galaxies.

**Key words:** methods: data analysis — techniques: spectroscopic — stars: general — galaxies: stellar content

## 1 INTRODUCTION

The Large Sky Area Multi-Object Spectroscopic Telescope (LAMOST), now being built in China, is a huge sky survey program where more than $10^7$ spectra of faint celestial objects

* Supported by "863" National High Technology R&D program.

down to 20.5 mag will be collected. Most of the objects are galaxies and AGNs, but some stars will be accidentally included. Automatically separating stars from galaxies and AGNs via their spectra is a key issue in the automated classification of point and point-like sources using the PSF test of images. In the program not only the object type but the associated redshift value also needs to be determined by some automated methods. Moreover, the SNR of those spectra is expected to be very low, which further aggravated the problem of automatic classification.

To our knowledge, only automated classification of spectra with known redshifts is reported in the literature (Connolly et al. 1995; Galaz & Lapparent 1998; Zaritsky et al. 1995). How to automatically separate spectra with unknown redshifts seems a more difficult and challenging task.

In terms of spectra, stars and normal galaxies can be regarded as non-active objects, while active galaxies and AGNs as active objects. Then, the separation of stars from galaxies and AGNs can be divided into two stages, as shown in Fig. 1. The first stage separates the active and non-active objects; the second stage, the stars and the normal galaxies. Since the first stage has been reported in our previous paper, where the non-active objects and active objects were successfully distinguished by a PCA+SVMs method, our present work is concentrated on the second stage, i.e., how to effectively extract stars from the non-active objects obtained from the first stage. In this work, a method based on PCA and SMM-RBFNN is used.

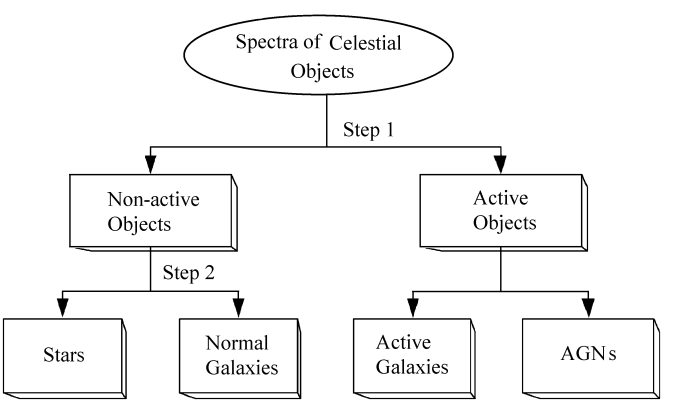


Fig. 1 Flow of spectral classification of celestial objects.

Stars and normal galaxies both have their own benchmark classifications. For example, the MK classification of stellar spectra classifies stars into seven main spectral types in the order of decreasing temperature, namely: O, B, A, F, G, K, M. Each type is in turn subdivided into 10 subtypes, ranging from 0 (the hottest) to 9 (the coolest), i.e., A0, A1, A2, ... A8, A9. Stellar spectra of the seven main types are shown in Fig. 2. Galaxies are classified into three broad groups based on their morphological structure according to the famous Hubble classification. They are ellipticals E, spirals S and irregulars I. In the present paper, we summarize normal galaxies spectra into four types, E0, S0, Sa, and Sb, which are shown in Fig. 3. For a normal galaxy, as a result of its stellar composition, its spectrum is quite similar to stellar spectra of F, G and K type, this contributing to the difficulty of classification.

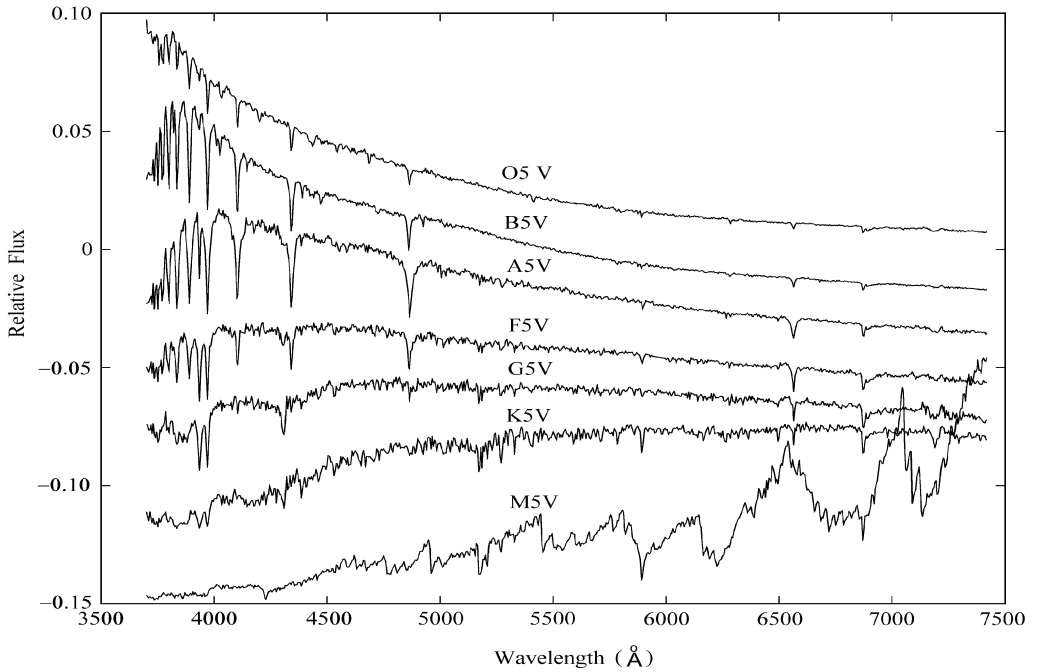


Fig. 2 Seven main types of stellar spectra.

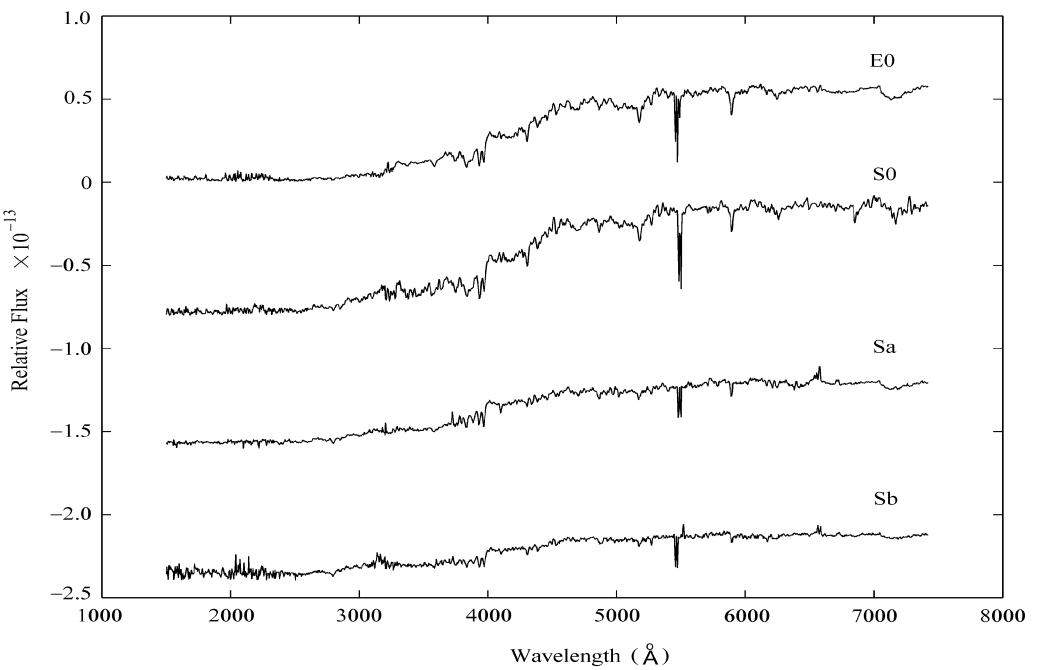


Fig. 3 Four templates of normal galaxies.

The organization of the present paper is as follows: In Section 2, we introduce the data set used in our experiments; In Section 3, the reduction of spectra dimension by a PCA method is described. In Section 4, spectral classification with SMM-RBFNN is elaborated. The experimental results are reported in Section 5, and finally some concluding remarks are given in Section 6.

## 2 THE DATA SET

### 2.1 Stellar Spectra

The stellar spectra used in our experiments are mostly selected from Astronomical Data Center (ADC). Among them, 161 stellar spectra given by Jacoby et al. (1984) are used in our training set and 131 observed by Pickles et al. (1998) are used in our testing set as well as 96 by Pickles et al. (1985). The spectra taken from the above libraries have respective resolutions of 1.4, 5 and 12 Å. In addition, we also used noise-added spectra with empirically controlled SNR by adding Gaussian noises.

### 2.2 Simulated Galaxy Spectra

Normal galaxy spectra in our training set were simulated from four templates of the types E, S0, Sa and Sb, in the galaxies spectrum atlas discussed by Kinney et al. (1996) and Calzetti et al. (1994). The templates are plotted in Fig. 3. The simulated spectra are generated from the above rest-frame templates by changing the redshift according to the equation

$$\lambda = (1 + z)\lambda_0. \tag{1}$$

Since the rest-frame templates cover the wavelength range of 1200–10 000Å, we could obtain 240 simulated redshifted normal galaxy spectra with redshift values ranging from 0.01–1.2 with a step of 0.02. In the testing set, we use 861 observed spectra selected from SDSS with average SNR of 13. Furthermore, the simulated "noise-added" redshifted spectra are used to test the robustness of our classifier.

All spectra are digitized and linearly interpolated to the wavelength range of 3800–7420 Å with a step of 5 Å.

## 3 SPECTRAL DIMENSION REDUCTION WITH PCA

Principal Components Analysis (PCA) (Bian & Zhang 1999; Huang et al. 2000; Qin et al. 2001) is a good tool for dimension reduction, data compression and feature extraction. In our work, we use PCA to remove redundant or insignificant information of spectral data to obtain a reduced dimension input space for our classifier. First, all the spectra, $\boldsymbol{X'}_{N\times M}$, are normalized by

$$\boldsymbol{X}_{N\times M} = (x_{ij})_{N\times M} = \left(\frac{x'_{ij} - \overline{X'_j}}{S'_j}\right)_{N\times M}, \qquad (i = 1, \ldots N; j = 1, \ldots M), \tag{2}$$

where $\boldsymbol{X}_{N\times M}$ is the normalized spectra matrix, $\overline{X'_j}$ and $S'_j$ the average and the standard bias of the $j$th column of $\boldsymbol{X'}$. Then the Principal Components (PCs) of stellar and normal galaxy spectra, $\boldsymbol{P}_{M\times M}$, can be obtained by calculating the eigenvectors of the covariance matrix of the spectra matrix. In Fig. 4 the eigenvalues are plotted in decreasing order. Since a normalized

spectrum (say $\boldsymbol{X}_{1\times M}$) could be represented by the principal components in the form:

$$\boldsymbol{X}_{1\times M} = \sum_{i=1}^{M} a_i \boldsymbol{P}_{i\times M}, \tag{3}$$

where, $a_i$ is the weight of spectrum $\boldsymbol{X}_{1\times M}$ for the $i$th PC. In Fig. 5 the error of spectral reconstruction with the reduced number of PCs is plotted. From the curves in Figs. 4 and 5, we can find the reduced number of PCs that could reconstruct the spectrum within tolerable limits. Since the first 50 PCs just have 1.48% reconstruction error, we choose them to define a 50-dimensional subspace and map spectra to be separated on it to obtain 50-dimensional vectors.

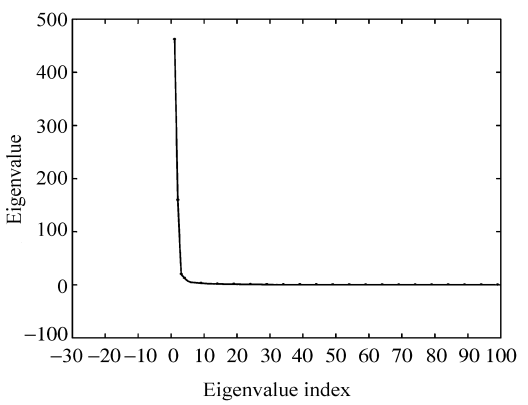


Fig. 4 Eigenvalue in decreasing order.

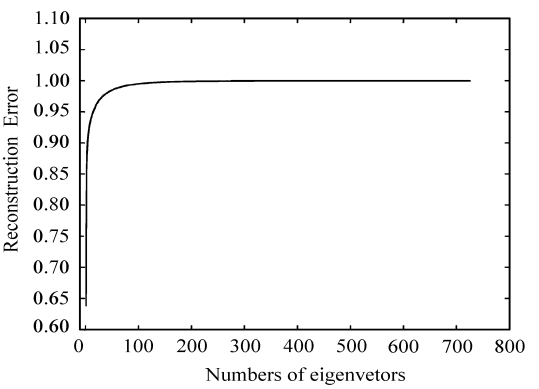


Fig. 5 Spectral reconstruction error with reduced number of PCs.

## 4 SPECTRAL CLASSIFICATION WITH SMM-RBFNN

### 4.1 SMM-RBFNN

Input spectra are projected to the subspace defined by the $q$ PCs ($q = 50$ in our case), and then SMM-RBFNN is invoked for the final classification. As we know, BPNN (Gupta et al. 1994; Coryn et al. 1998) has already had many applications on astronomical data reduction. However, there are two main drawbacks in BPNN, one is its slow convergence rate, and the other is that it sometimes falls into local minimum instead of global minimum. The slow convergence rate will pose a computational problem, and local minima will degrade the classification accuracy. To acquire better performances in both the training time and the recognition accuracy, we propose a new method called SMM-RBFNN. In SMM-RBFNN, we suppose the conditional probability function in our system as a mixture modeling (Webb 1999) with the following form:

$$p(\boldsymbol{z}|\boldsymbol{x}, \Theta) = \sum_{j=1}^{H} P(j) p(\boldsymbol{z}|\boldsymbol{x}, \Theta_j) \ , \tag{4}$$

$\boldsymbol{x}$ and $\boldsymbol{z}$ being the input and output vectors, (dimensions $d_x$ and $d_z$) and $\Theta$ the parameter set. $H$ is the number of mixture components and $P(j)$ is the $j$th weight of $j$th component, subject

to $P(j) \geq 0, \sum_{j=1}^{H} P(j) = 1$. Each component is assumed to be a Gaussian:

$$p(\boldsymbol{z}|\boldsymbol{x}, \Theta_j) = G(\boldsymbol{z}, \boldsymbol{g}_j(\boldsymbol{x}, \mathbf{W}), \sigma_j^2) \ , \tag{5}$$

$$G(\boldsymbol{z}, \boldsymbol{g}_j(\boldsymbol{x}, \boldsymbol{W}), \sigma_j^2) = \frac{1}{(2\pi\sigma_j^2)^{d_z/2}} \exp\left\{-\frac{1}{2\sigma_j^2}||\boldsymbol{z} - \boldsymbol{g}_j(\boldsymbol{x}, \mathbf{W})||^2\right\}, \tag{6}$$

where the map function $\boldsymbol{g}_j(\boldsymbol{x}, \mathbf{W})$ is assumed the RBFNN, whose components are written as:

$$\begin{aligned} g_{j,l}(\boldsymbol{x}, \mathbf{W}) &= \sum_{i=1}^{K_j} w_{l,i}\phi_i(||\boldsymbol{x} - \mu_i||) + w_{l,0} \\ &= \sum_{i=0}^{K_j} w_{l,i}\phi_i(||\boldsymbol{x} - \mu_i||), \quad l = 1, \cdots d_z \ , \end{aligned} \tag{7}$$

where $w_{(l,0)}$ is a set of bias constants, $\phi_0(||\boldsymbol{x} - \mu_i||) \equiv 1$ and

$$\phi_i(||\boldsymbol{x} - \mu_i||) = \exp\Big[-\frac{1}{2\gamma_i^2}||\boldsymbol{x} - \mu_i||^2\Big]. \tag{8}$$

In the present paper, all RBF are assumed to be Gaussian radial basis functions. From the above formulation, our neural network is composed of $H$ RBFNNs, and the $j$th RBFNN includes $k_j$ RBF with the same or different center $\mu_i$ and variance $\gamma_i$. $\mathbf{W}$ denotes the weight matrix in RBFNN. Given the learning samples $D = \{\boldsymbol{x}_n, \boldsymbol{z}_n\}_{n=1}^N$, the cost function of system, which also called the error function, can be written as:

$$E = -\ell(\Theta) = -\sum_{n=1}^{N} \ln p(\boldsymbol{z}_n|\boldsymbol{x}_n, \Theta) = -\sum_{n=1}^{N} \ln \sum_{j=1}^{H} p(j)p(\boldsymbol{z}_n|\boldsymbol{x}_n, \Theta_j), \tag{9}$$

where $\ell(\boldsymbol{\Theta})$ is the log-likelihood function of system, $\boldsymbol{\Theta}$ is the parameter set $\{\sigma, \mathbf{W}, \gamma, \mu\}$.

### 4.2 EM-Like Learning Algorithm in SMM-RBFNN

Given the learning samples, the parameters in SMM-RBFNN classifier can be estimated by maximizing the log-likelihood function. From Eq. (9), EM-like learning algorithm (Webb 1999; Guo & Xu 1999) could be derived with matrix differential calculus method (Magnus & Neudecker 1999), which is described as the following steps:

**E-step:**

Fix up the parameter $P(j)^{\text{old}}$ and $\boldsymbol{\Theta}^{\text{old}}$, compute the variable $h(j, n)$ according to the following equation:

$$h(j, n) = \frac{P(j)^{\text{old}} G(\boldsymbol{z}_n, \boldsymbol{g}_j(\boldsymbol{x}_n, \mathbf{W}^{\text{old}}), (\sigma_j^2)^{\text{old}})}{\sum_{i=1}^{H} P(i)^{\text{old}} G(\boldsymbol{z}_n, \boldsymbol{g}_i(\boldsymbol{x}_n, \mathbf{W}^{\text{old}}), (\sigma_i^2)^{\text{old}})}. \tag{10}$$

**M-step:**

Compute the new estimates of the parameters $P(j)^{\text{new}}$ and $\boldsymbol{\Theta}^{\text{new}}$ according to the following formulae:

$$P(j)^{\text{new}} = \frac{1}{N}\sum_{n=1}^{N} h(j, n), \tag{11}$$

$$\boldsymbol{\Theta}^{\text{new}} = \arg\max\{\ell(\boldsymbol{\Theta})\}. \tag{12}$$

Equation (12) can also be written as

$$\boldsymbol{\Theta}_j^{\text{new}} = \boldsymbol{\Theta}_j^{\text{old}} + \eta \nabla \ell(\boldsymbol{\Theta})|_{\boldsymbol{\Theta}=\boldsymbol{\Theta}^{\text{old}}}, \tag{13}$$

which shows their $\boldsymbol{\Theta}^{\text{new}}$ is a correction of $\boldsymbol{\Theta}^{\text{old}}$ along the gradients by a factor of $\eta \nabla \ell(\boldsymbol{\Theta})$. More specifically, Eq. (13) can be expressed in term of each component $\{\sigma, \mathbf{W}, \gamma, \mu\}$ as

$$(\sigma_j^2)^{\text{new}} = (\sigma_j^2)^{\text{old}} + \eta \frac{\partial \ell(\boldsymbol{\Theta})}{\partial \sigma_j^2}|_{\boldsymbol{\Theta}=\boldsymbol{\Theta}^{\text{old}}}, \tag{14}$$

where $\eta$ is a learning factor. The partial derivative of $\ell(\boldsymbol{\Theta})$ with respect to each parameter can be obtained as

$$\frac{\partial \ell(\boldsymbol{\Theta})}{\partial \sigma_j^2} = \sum_{n=1}^{N} h(j,n) \left[ -\frac{d_z}{2\sigma_j^2} + \frac{\left\| \boldsymbol{z}_n - \boldsymbol{g}_j(\boldsymbol{x}_n, \mathbf{W}_j) \right\|^2}{2(\sigma_j^2)^2} \right], \tag{15}$$

$$\frac{\partial \ell(\boldsymbol{\Theta})}{\partial (w_{i,l})_j} = \sum_{n=1}^{N} h(j,n) \left[ \frac{z_{n,l} - g_{j,l}(\boldsymbol{x}_n, \mathbf{W})}{\sigma_j^2} \phi_i(||\boldsymbol{x}_n - \mu_i||) \right], \tag{16}$$

$$\frac{\partial \ell(\boldsymbol{\Theta})}{\partial (\mu_{i,m})_j} = \sum_{n=1}^{N} h(j,n) \left\{ \frac{1}{2\sigma_j^2} [\boldsymbol{z}_n - \boldsymbol{g}_j(\boldsymbol{x}_n, \mathbf{W}_j)]^T (\boldsymbol{w}_i)_j \phi_i(||\boldsymbol{x}_n - \mu_i||) \left[ \frac{x_{n,m} - \mu_{i,m}}{2\gamma_i^2} \right] \right\}, \tag{17}$$

$$\frac{\partial \ell(\boldsymbol{\Theta})}{\partial (\gamma_i^2)_j} = \sum_{n=1}^{N} h(j,n) \left\{ \frac{1}{2\sigma_j^2} [\boldsymbol{z}_n - \boldsymbol{g}_j(\boldsymbol{x}_n, \mathbf{W}_j)]^T (\boldsymbol{w}_i)_j \phi_i(||\boldsymbol{x}_n - \mu_i||) \left[ \frac{||\boldsymbol{x}_n - \mu_i||^2}{2(\gamma_i^2)^2} \right] \right\}, \tag{18}$$

where the subscript $j$ denotes the parameters in $j$th mixture component and $\boldsymbol{w}_i = (w_{i,1}, w_{i,2}, \cdots w_{i,d_z})$ is the vector in $i$th row of the weight matrix.

Applying the above EM-like algorithm to estimate the parameter in SMM-RBFNN, we then build the classifier for separating stars from normal galaxies by their spectra.

### 4.3 Spectral Classification via SMM-RBFNN

We investigate the SMM-RBFNN for classifying stellar spectra from normal galaxies with the architecture depicted in Fig. 6. The SMM-RBFNN consists of $H$ RBFNNs of each is regarded as one component in the statistical mixture modelling. Figure 7 shows the $j$th RBFNN architecture with $K_j$ RBF nodes in the hidden layer, where $j = 1, 2, \ldots, H$ and $l = 1, 2, \ldots, d_z$. The input vector $\boldsymbol{x}$ is 50-dimensional, obtained by PCA and output $\boldsymbol{z}$ is 2-dimensional, denoting the probabilities belonging to stars or normal galaxies, that are, $d_x = 50$ and $d_z = 2$.

With different numbers of RBF neurons, the RBFNN performs differently in the classification of input samples and takes different statistical contribution to the final classifier. Therefore, through the statistical mixture modeling with the classification results from $H$ RBFNNs, we could obtain a better classifier in separating stars from normal galaxies with unknown redshifts.

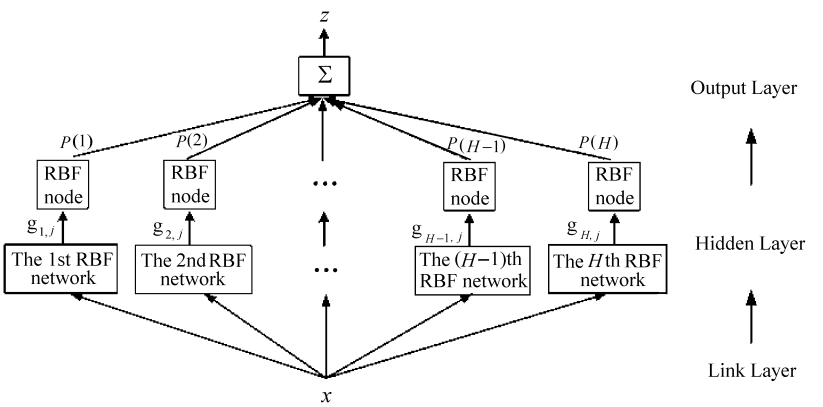


Fig. 6 Architecture of SMM-RBFNN.

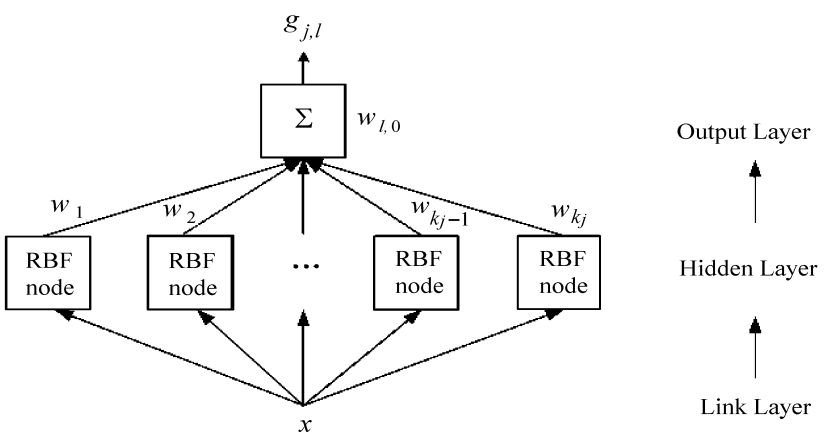


Fig. 7 $j$th RBFNN architecture with $K_j$ RBF nodes ($j = 1 \ldots H, l = 1 \ldots d_z$).

## 5 EXPERIMENTS AND RESULTS

In our experiments, the training data set for the SMM-RBFNN classifier includes 161 stellar spectra and 240 simulated redshifted spectra of normal galaxies, as described in Sections 2.1 and 2.2. We use five RBFNNs in SMM-RBFNN whose RBF nodes are respectively set to 150, 200, 250, 300 and 360.

### 5.1 Initialization in SMM-RBFNN

The parameters in five RBFNNs include the variances $\gamma_{j,i}^2$ and centers $\mu_{j,i} (j = 1 \ldots 5, i = 1 \ldots K_j)$. We set $\gamma_{j,i}^2 = 1$ for all $i$, $j$, and obtain the $\mu_{j,i}$ by pre-clustering with K-means algorithm (Bian & Zhang 1999). Then, the initial probabilities $P(j)$ of the five RBFNNs in

SMM are set to the average value of 0.2. The weight matrix $\mathbf{W}^j$ in the $j$th RBFNN is initialized by

$$\mathbf{W}^j_{K_j\times 2} = (\Phi^{-1}_{K_j\times 401})\boldsymbol{Z}_{401\times 2}, \tag{19}$$

where $\Phi^j$ is the matrix of RBF nodes output in the $j$th RBFNN and $\boldsymbol{Z}$ is the expectation of final output. The variance $\sigma_j^2$ can be set by

$$\sigma_j^2 = \frac{1}{401\times 2}\sum_{n=1}^{401}||Z_n - \sum_{j=1}^{5}P(j)g_j(x_n,\mathbf{W}^j)||^2 . \tag{20}$$

### 5.2 Training and Testing Results

After setting the initial values in SMM-RBFNN, we apply the EM-like learning algorithm to train the networks. The detailed steps are shown in Section 4.2. In our experiments the training procedure terminates and reaches the extremum just after several loops and takes less than a few seconds. The training results are partly shown in Table 1.

**Table 1** A Part of Training Parameter Results of SMM-RBFNN

| Parameter | $j=1$ | $j=2$ | $j=3$ | $j=4$ | $j=5$ |
|---|---|---|---|---|---|
| $P(j)$ | 0.0345 | 0.0849 | 0.1984 | 0.3573 | 0.8624 |
| $\sigma_j^2$ | 0.0206 | 0.0206 | 0.0206 | 0.0206 | 0.0207 |

We used the test set with 227 stellar spectra and 861 SDSS spectra of normal galaxies with average S/N ratio= 13, to test our classifier and obtained a correct recognition rate of 98.68% and 96.52%. Using spectra with different S/N ratio, which are generated by adding Gaussian noises on the standard spectra as described in Sections 2.1 and 2.2, the correct recognition rate vs. SNR is plotted in Figs. 8 and 9. The curve in Fig. 8 is based on 388 noise-added stellar spectra and that in Fig. 9 is based on 240 noise-added spectra of normal galaxies. The results show that much higher correct recognition rates are obtained with the SMM-RFBNN.

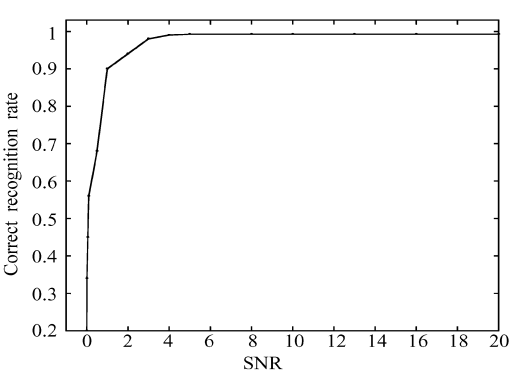


Fig. 8 Correct recognition rate vs. SNR (noise-added stellar spectra).

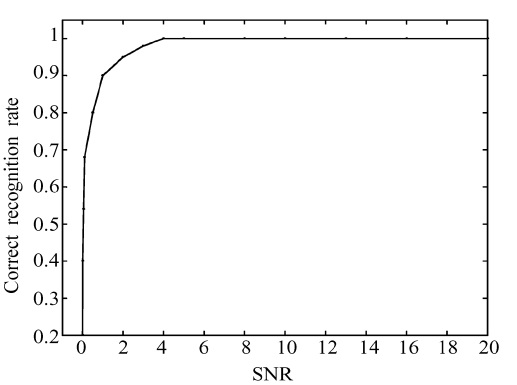


Fig. 9 Correct recognition rate vs SNR (noise-added spectra of normal galaxies).

## 6 CONCLUSIONS

Usually point sources can be separated from extended sources by the PSF test. However, point-like sources are a possible mixture of galaxies, AGNs and stars. Therefore, the question of separating stars from galaxies and AGNs by spectra automatically is not well settled. We propose a new automated classification technique to separate stars from normal galaxies with unknown redshift values via their spectra. By combining the method in this paper and in our previous one about automated separation of non-active objects and active objects via spectra, a complete solution of separating stars from galaxies or AGNs has been obtained.

In our proposed technique, first the PCA is adopted to reduce the dimension of spectra to 50 with a reconstruction error of 1.48%, then, the reduced inputs are feded to SMM-RBFNN to separate stars from normal galaxies. By simulating redshifted spectra of normal galaxies, we train the SMM-RBFNN classifier on a training set with different types and a large range of redshift values. Since RBF parameter initialization using the K-means can approach the optimal value, the EM-like learning algorithm can be applied in SMM-RBFNN to quickly reach the extremum just after a few loops. Meanwhile, it is not necessary to use large numbers of training data to reach the extremum of SMM-RBFNN. Experiments have shown that our method can obtain a correct classification rate as high as 99.22% for standard stellar spectra, and 96.52% for the SDSS spectra of normal galaxies with average SNR of 13. The experimental results also show that our technique is robust and computationally efficient. Therefore it is suitable for automated classification of voluminous spectra with low SNR from sky surveys.

**Acknowledgements** This work was supported by "863" National High Technology R&D program (No.2001AA133010) and the LAMOST project of National Astronomical Observatories, Chinese Academy of Sciences.

### References

Connolly A. J., Szalay A. S., Bershady M. A., Kinney A. L., Calzetti D.,1995, AJ, 110, 1071
Galaz G., Lapparent V., 1998, A&A, 332, 459
Zaritsky D., Zabludoff A. I., Jeffrey A. W., 1995, AJ, 110, 1602
Jacoby G. H., Hunter D. A., Christian C. A., 1984, ApJS, 56, 257
Pickles A. J., 1998, PASP, 110, 863
Pickles A. J., 1985, ApJS, 59, 33
Kinney A. L., Calzetti D., Bohlin R. C., McQuade K., Storchi-Bergmann T., 1996, AJ, 467, 38
Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, AJ, 429, 582
Kent S. M., 1994, Ap&SS, 217, 27
Bian Z. Q., .Zhang X. G., 1999, Pattern Recognition, 2nd ed., Beijing: Tsing Hua University Press
Huang L. Y., Sun F. M., Hu Z. Y., 2000, In: A. K Jain, J. Kittler, eds., ICPR'2000, Vol.2, 499
Qin D. M., Hu Z. Y., Zhao Y. H., 2001, In: J. Shen, S. Pankanti, R. Wang, eds., Proc. SPIE, Object Detection, classification, and Tracking technologies, 4554, 268
Gupta R. A., Gulati R., Gothoskar P., Khobragade S., 1994, AJ, 426, 340
Coryn A. L., Bailer-Jones C. A. L., Irwin M., Hippel T. V., 1998, MNRAS, 298, 361
Webb A., 1999, Statistical Pattern Recognition, London: Oxford University Press
Guo P., Xu L., 1999, In: S. Verlag, ed., Proc. ICONIP, Vol. I, Perth: ICONIP, 426
Magnus J. R., Neudecker H., 1999, Matrix Differential Calculus with applications in statistics and econometrics, New York: John Wiley & Sons