

# Remote Sensing Image classification based on Optimized Support Vector Machine

Liqian cheng, Wenxing Bao\*

School of Computer Science and Engineering, Beifang University of Nationalities,  
No. 204 Wenchang, North-Street, Xixia District, Yinchuan Ningxia, P. R. China 750021

\*Corresponding author, e-mail: [bw71@163.com](mailto:bw71@163.com)

## Abstract

To resolve the problem of wetland remote sensing image classification, this paper presents an improved classification algorithm. In this algorithm, genetic algorithm (GA) selection, crossover operation is introduced to the standard particle swarm optimization algorithm (PSO) to form a hybrid particle swarm optimization algorithm (GAPSO). The hybrid algorithm can exploit the advantages of the genetic algorithm and particle swarm algorithm respectively to the full to obtain the global optimal parameters of support vector machine (SVM). Thus the wetland remote sensing image can be classified more accurately. Taking Ningxia Shahu wetland remote sensing images as an example, this paper makes a classification of wetland remote sensing images using optimized support vector machine, and the outcome of the experiment shows that this algorithm has better classification effect than that of similar algorithms.

**Keywords:** remote sensing image classification, support vector machine (SVM), parameter optimization, hybrid particle swarm optimization (GAPSO)

**Copyright © 2014 Institute of Advanced Engineering and Science. All rights reserved.**

## 1. Introduction

Wetland is one of the most important natural resources and the ecological system, which plays an important role in climate regulation, water conservation, environmental purification, dispersion of flood and the conservation of biological diversity. The rapid development of remote sensing technology has made it become an important means of wetland monitoring and classification, and has been widely used in the monitoring of wetlands [1].

SVM (Support Vector Machine), a learning approach originally developed by Vapnik and co-workers according to the statistical theory is based on the structural risk minimization principle. It can solve the problems of small sample, nonlinear, high dimension commendably and become one of the hotspot of machine learning. It has been applied to solve classification problems, function approximation and time series prediction successfully etc. In recent years, an increasing number of scholars use SVM to image classification and information extraction and SVM has often been found to provide higher classification accuracies than other widely used pattern recognition techniques, such as the maximum likelihood and the multilayer perceptron neural network classifiers. Furthermore, SVM appears to be especially advantageous in the presence of heterogeneous classes for which only few training samples are available [2-4]. Like other learning algorithms, the function of SVM heavily depends on the parameters of its model. The SVM classifier trained by different parameters varies widely in performance. If the value of the parameter is improper, then SVM could not achieve the learning effect as it is expected. Before training the SVM classifier it is necessary for us to optimize the parameters to obtain the best combination of parameters. There are several frequently-used methods such as grid search algorithm, genetic algorithm (GA), gradient descent algorithm etc. The grid search algorithm is intuitive and easy to realize and is a good choice when there is few parameters, however, when applying grid search method, one might need to increase the parameter range and/or decrease the step size to increase the accuracy of the optimal solution. However, this will result in a cumbersome time consuming search process [5]. Gradient descent method has improved significantly in the calculation of time, but the stability is not high [6].

Particle swarm optimization (PSO) is put forward by Eberhart and Kennedy in 1995, which was inspired by the social behavior of bird flocking and fish schooling [7]. Its main feature is simple in principle, fewer parameters, fast convergence and required little domain

knowledge [8]. This algorithm has obtained widespread application in fields of the function optimization, neural network training, and can be achieved good results. Just as other evolutionary algorithms, it is easy to fall into local minima and premature convergence. Thus, Lovbjerg [9] proposed a hybrid particle swarm algorithm which combines the idea of genetic algorithm and particle swarm optimization algorithm. Genetic algorithms has been sucessfully used to optimaze inelligent algorithms such as artificial neural network [10]. The combination of the two methods can play to their strengths on the premise of guaranteeing the diversity of particle swarm, in addition, it has accelerated the speed of convergence and enhanced the ability to jump out of local minima and find global optimal solution.

This paper uses SVM classifier optimized by hybrid particle swarm algorithm (GAPSO\_SVM) to classify the remote sensing image of Ningxia Shahu, it also applies grid search algorithm, genetic algorithm and particle swarm optimization algorithm to optimize SVM in order to compare with GAPSO\_SVM. The classification result shows that the classification accuracy of GAPSO\_SVM is better than other three classifiers.

The rest of this paper is organized as follows. Section 2 gives a brief introduction to SVM, the importance of the parameters of SVM, PSO algorithm and describes how the proposed approach optimizes parameters. Section 3 detailed experiments are carried out to evaluate Grid\_SVM, GA\_SVM, PSO\_SVM, and our GAPSO\_SVM method. Finally, the conclusions are presented in Section 4.

## 2. Research Methods

### 2.1. SVM

SVM is proposed to solve linearly separable problem originally. Let us assume that the training set consists of  $n$  vectors  $\{(x_i, y_i), i=1,2,\dots,n\}$ . A target  $y_i \in (-1,1)$  is associated to each vector  $x_i$ . If the two classes are linearly separable, then at least one hyperplane  $\omega \cdot x_i + b = 0$  can be found to separate the two classes without errors. SVM classifier will look for an optimal hyperplane that maximizes the distance between the closest training sample and the separating hyperplane. Normalizing  $\omega$  and  $b$ , the distance between the two types of the sample points is  $\frac{2}{\|\omega\|}$ , making  $\frac{2}{\|\omega\|}$  get the maximum means making  $\frac{\|\omega\|^2}{2}$  get the minimum, all above can be described by the mathematical language, namely:

$$\begin{cases} \min \frac{\|\omega\|^2}{2} \\ s.t. y_i(\omega \cdot x_i + b) \geq 1, i=1,2,\dots,n \end{cases} \quad (1)$$

The problem can be obtained by solving saddle points of LaGrange function.

$$\phi(\omega, b, \alpha_i) = \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^n \alpha_i [y_i(\omega \cdot x_i + b) - 1] \quad (2)$$

Where  $\alpha_i \geq 0$  are the Lagrange multipliers [11].

In practical applications, most of the problem is nonlinear, and then linearly separable SVM is incapable of action. For the nonlinearly separable problem, the method usually used is nonlinear mapping  $\phi: \mathbb{R}^d \rightarrow \mathbb{H}$  that is the sample of the original input is mapped to a higher dimensional space. Then construct optimal separating hyperplane in the high dimensional feature space. According to the functional theory, as long as one kind of kernel function  $K(x_i, x_j)$  satisfies Mercer condition, it corresponds to an inner product of the inner space. Therefore, linear classification can be achieved by using an appropriate inner product function

$K(x_i, x_j)$  in the optimal hyperplane after nonlinear transformation. After mapping into a high dimensional feature space the corresponding dual problem becomes:

$$\begin{cases} \max Q(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, x_j) \\ \text{s.t.} \begin{cases} \sum_{i=1}^n \alpha_i y_i = 0 \\ 0 \leq \alpha_i \leq C \end{cases}, i = 1, 2, \dots, n \end{cases} \quad (3)$$

Thus the optimal classification function is:

$$f(x) = \text{sgn}\left(\sum_{i=1}^n \alpha_i^* y_i K(x_i, x_j) + b^*\right) \quad (4)$$

In classification of SVM, an appropriate kernel function and model parameters has a great influence on the performance of classification algorithms. Research shows that RBF kernel function performs better in classification accuracy and classification time than other kinds of kernel function in the classification of remote sensing images [12, 13]. Therefore, RBF kernel function is employed for the classification of TM images in this study. Wherein, the model parameters including kernel function parameters and the error penalty factor. In the SVM training process, the error penalty factor C (ie, upper bound of support vector coefficient  $\alpha$ ) is to achieve a compromise between the proportion of misclassified samples and the complexity of the algorithm. If the value of the error penalty factor C is too low, which means that the punishment for the empirical error is little; the complexity of the machine learning experience is small and the empirical risk value is large, which means "learning deficiency"; if C is infinite, all constraints must be met, which means that the training sample must be accurately classified, i.e. "over learning." The performance of SVM is also directly affected by the pros and cons of nuclear parameters. As the kernel functions, mapping functions as well as high-dimensional feature space have a one to one correspondence. As soon as the kernel function is determined, it implicitly determines the mapping function and high dimensional feature space. Kernel parameter changes actually alter the mapping function implicitly, thus changing the complexity of the sample in the distribution of high-dimensional feature space [Error! Bookmark not defined.].

## 2.2. Particle Swarm Optimization

In particle swarm optimization, each individual is called a "particle", which represents a potential solution. Assuming  $X_i = (X_{i1}, X_{i2}, \dots, X_{id})$  is the position of the i-th particle in D-dimensional space; according to the preset fitness function (related to the problem) we calculate the current fitness value of  $X_i$  and evaluate the quality of the particle positions.  $V_i = (V_{i1}, V_{i2}, \dots, V_{id})$  is the flight speed,  $pbest_i = (p_{i1}, p_{i2}, \dots, p_{id})$  represents the optimal position for the i-th particle so far,  $gbest = (g_1, g_2, \dots, g_d)$  represents the optimal position for the whole particles so far. In each iterations, the particles update velocity and position according to the following formula:

$$\begin{aligned} v_{id}^{k+1} &= \omega v_{id}^k + c_1 r_1 (p_{id}^k - x_{id}^k) + c_2 r_2 (g_{id} - x_{id}^k) \\ X_{id}^{k+1} &= X_{id}^k + v_{id}^{k+1} \end{aligned} \quad (5)$$

$i = 1, 2, \dots, m$ ,  $d = 1, 2, \dots, D$ ,  $\omega$  is the inertia weight, it plays a role of balancing the global search and local search,  $k$  is the iteration number,  $r_1$  and  $r_2$  are  $[0,1]$  random numbers, these two parameters are used to maintain the diversity of the population,  $c_1$  and  $c_2$  are learning factors which make particles have self-summary and the ability to learn from outstanding individuals of groups [14].

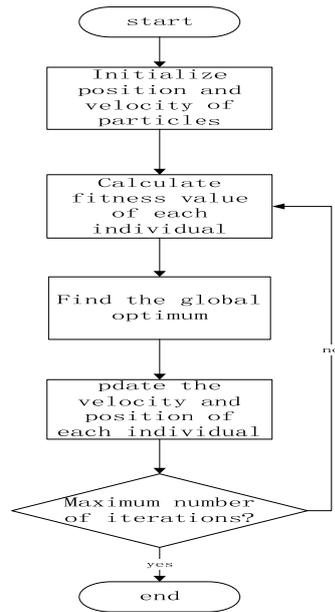


Figure 1. Flowchart of PSO Algorithm.

In hybrid particle swarm optimization algorithm, particles are endowed with a crossover probability; the probability of this hybridization is user-determined and independent of the particle's fitness value. In each iterations, selecting a specified number of hybrid particles based on the crossover probability and putting them in a pool, particles in the pool randomly crossed to produce the same number of offspring and parent particles are replaced with progeny particles, which can keep the number of particles constant. The nature of the genetic algorithm can improve the diversity of particle and particle search ergodicity but also to prevent particle swarm into local optimum and improve the algorithm against premature to better optimize the SVM classifier.

The position and velocity of the offspring is updated as follows:

$$\text{child}_1 x_i = p_i \text{parent}_1(x_i) + (1.0 - p_i) \text{parent}_2(x_i) \quad (6)$$

$$\text{child}_2 x_i = p_i \text{parent}_2(x_i) + (1.0 - p_i) \text{parent}_1(x_i) \quad (7)$$

$$\text{child}_1(v_i) = \frac{\text{parent}_1(v_i) + \text{parent}_2(v_i)}{|\text{parent}_1(v_i) + \text{parent}_2(v_i)|} * |\text{parent}_1(v_i)| \quad (8)$$

$$\text{child}_2(v_i) = \frac{\text{parent}_1(v_i) + \text{parent}_2(v_i)}{|\text{parent}_1(v_i) + \text{parent}_2(v_i)|} * |\text{parent}_2(v_i)| \quad (9)$$

Since SVM based on RBF kernel function has two parameters  $C$  and  $\gamma$ , each particle represents a group of SVM kernel parameters  $(C, \gamma)$ , and classification accuracy is selected to be the fitness function. The specific description of GAPSO\_SVM is as follows:

- 1) Set the size of the particle swarm, fitness error limits, maximum number of iterations allowed, inertia weight and learning factors  $c_1$ ,  $c_2$  and other parameters;
- 2) Initialize the position and velocity of particles;
- 3) Calculate the fitness of particles, set the current particles as the pbest, set the particle whose fitness is the best as the global optimum gbest;
- 4) Update the position and velocity of particles;

5) Select a specified number of particles based on the probability into the pool, the particles in the pool produce the same number of offsprings randomly by making hybrids with each other;

6) Check if it reaches the maximum number of iterations or the optimal solution; if it does then output the optimal solution otherwise proceed Step 3 to 6.

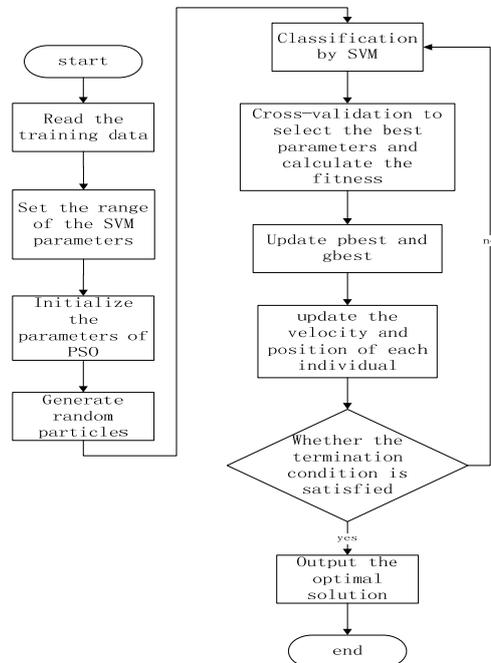


Figure 2. Flowchart of GAPSO\_SVM Model

### 3. Experiment and Analysis

The multispectral dataset used in our experiments is the Landsat TM5 remote sensing images of Ningxia Shahu wetland area taken on September 16, 2009.

The first step is to make the image preprocessing and feature extraction. This part is done with the help of the ENVI software. Spectral features are adopted, since the spectral feature is the basic feature of remote sensing images and using spectral characteristics in image classification is the most simple and effective method. The image is cut out in ENVI4.8 software and the size is 256\*256. Then the image is pretreated with PCA and the first three principal components which contain the most abundant characteristic information are chosen, the image as shown in Figure 3.

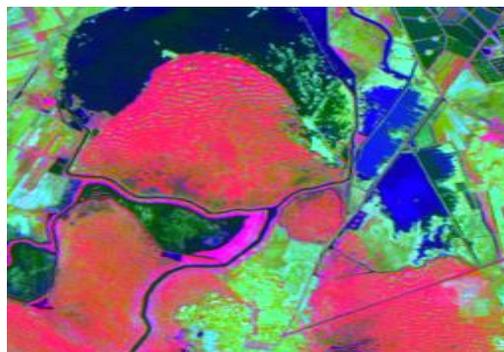


Figure 3. The Original Image

The second step is the realization of the algorithms and the image classification. SVM algorithm and its parameter optimization are implemented by using the LibSVM toolbox based on Matlab platform. LibSVM is mainly used for solving classification and regression problems. The toolkit also provides the source code, so that users can develop their own functions expediently. In both the standard PSO model and the hybrid model, the parameters are set as following. The searching range of SVM parameters  $C$  and  $\gamma$  are between 0.1 and 100. The cognition learning factor  $c1$  and the social learning factor  $c2$  are set to be 2, the inertia weight  $\omega$  is 0.8. The population size is set to be 20, and the number of iterations is set to be 10, 20 and 50 respectively. The crossover probability of GAPSO is set to be 0.9. The population size and the number of iterations of GA are the same as that of GAPSO, the crossover probability and mutation probability are 0.7 and 0.01, respectively. The searching bound of grid search method parameter  $C$ ,  $\gamma$  are set to be  $2^{-8}$  and  $2^8$ , and both of step values are set to be 0.8. To ensure the robustness experiment, the value of cross validation parameter  $k$  is set to be 5 folds.

The study area contains three classes: water, desert and vegetation. The numbers of training samples randomly selected from the study area are: 1664, 4459, and 1456, respectively. Then choose the whole image as test samples. Figure 4 shows the classification result when the iteration is set to be 50.

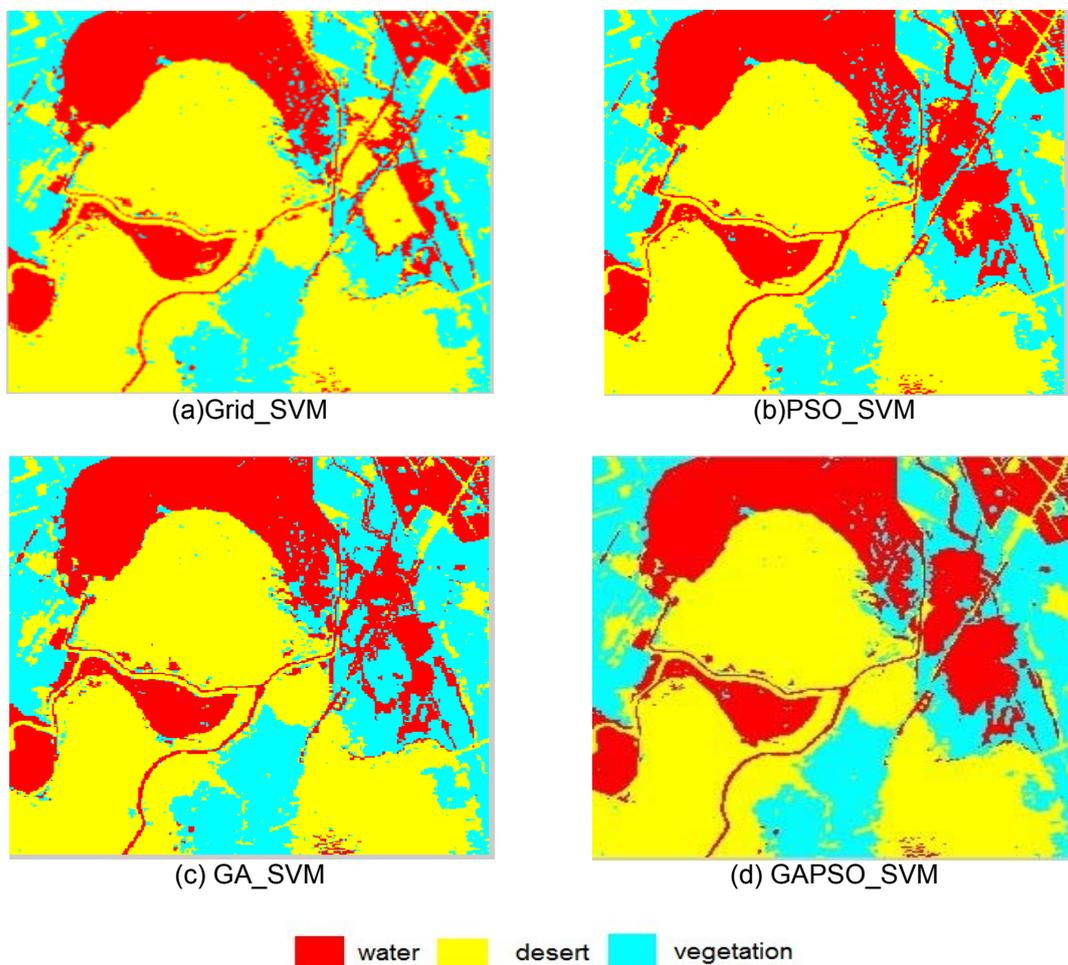


Figure 4. Classification Result when Iteration is 50

The comparison results obtained by the Grid\_SVM, GA\_SVM, PSO\_SVM and GAPSO\_SVM are shown in Table 1 to 3. Table 1 to 3 has shown the performance of the genetic algorithm, particle swarm algorithm and hybrid particle swarm optimization when the iterations

were 10, 20 and 50 respectively. The tables list the values of  $c$  and  $\gamma$ , C.V. accuracy, overall accuracy, Kappa coefficient and the time consumed.

Table 1. C.V. Accuracy, Overall Accuracy, Kappa Coefficient of 10 Generations

	Grid_SVM	PSO_SVM	GA_SVM	GAPSO_SVM
C.V.accuracy	99.5831	99.6278	99.6129	99.6426
$c/\gamma$	1/16	0.1/1.0740	2.1484/8.1542	86.3847/0.7884
overall accuracy	89.907%	99.259%	94.368%	99.328%
Kappa coefficient	0.82986	0.98785	0.90888	0.98899
Time(s)	1557.402	128.85	357	172.68

Table 2. C.V. Accuracy, Overall Accuracy, Kappa Coefficient of 20 Generations

	Grid_SVM	PSO_SVM	GA_SVM	GAPSO_SVM
C.V.accuracy	99.5831	99.6278	99.6129	99.6426
$c/\gamma$	1/16	0.1/1.0740	2.1484/8.1542	86.3847/0.7884
overall accuracy	89.907%	99.259%	94.368%	99.328%
Kappa coefficient	0.82986	0.98785	0.90888	0.98899
Time(s)	1557.402	241.47	463.12	223.28

Table 3. C.V. Accuracy, Overall Accuracy, Kappa Coefficient of 50 Generations

	Grid_SVM	PSO_SVM	GA_SVM	GAPSO_SVM
C.V.accuracy	99.5831	99.6278	99.6129	99.6426
$c/\gamma$	1/16	0.1/1.0740	2.1484/8.1542	86.3847/0.7884
overall accuracy	89.907%	99.259%	94.368%	99.328%
Kappa coefficient	0.82986	0.98785	0.90888	0.98899
Time(s)	1557.402	599.83	782.01	387.87

Figure 5 to 7 have shown the best fitness and average fitness of each generation of the PSO method, the GA and the GAPSO method with 10, 20 and 50 generations respectively.

Table 1 to 3 and Figure 4 illustrate that GAPSO\_SVM's performance of classification is the best compared with the other 3 models. Overall accuracy and Kappa coefficient are the highest of the 4 methods. GAPSO\_SVM's performance of classification is a little bit better than PSO\_SVM's. But with the increase in the number of iterations, the time that PSO\_SVM takes is longer than that GAPSO\_SVM takes. Grid\_SVM's performance of classification is the worst of the 4 methods and it takes the longest time as well. Figure 5 to 7 show that the GAPSO method achieves better results and has much faster convergence than both the GA and the PSO model. The average fitness of GAPSO is more stable and close to the best fitness.

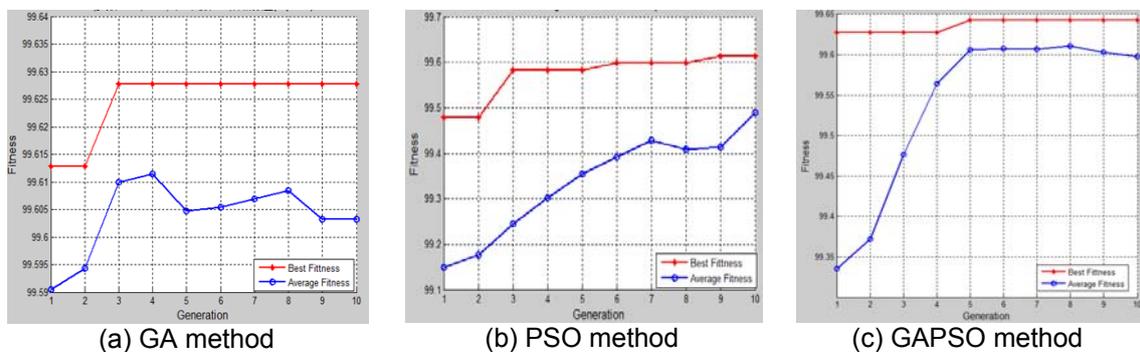


Figure 5. Fitness Curve of Three Methods with 10 Generations

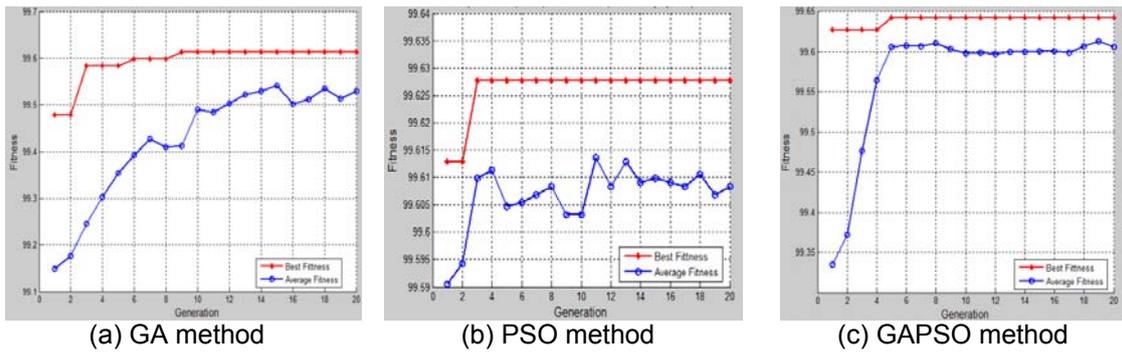


Figure 6. Fitness Curve of Three Methods with 20 Generations

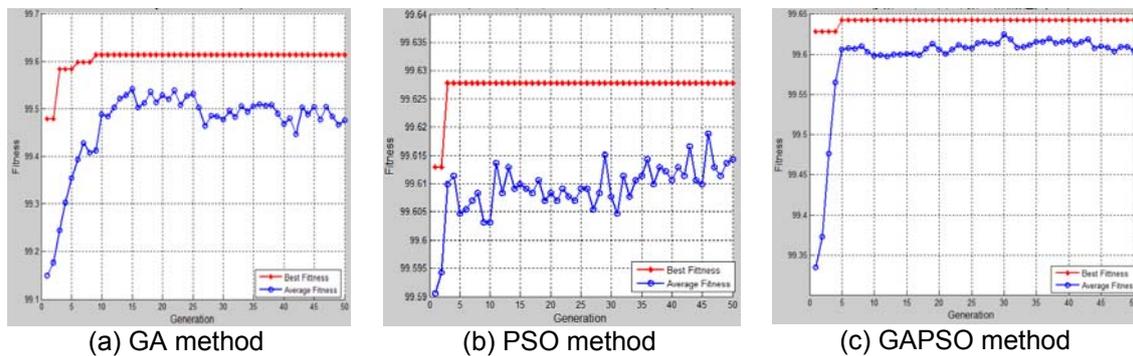


Figure 7. Fitness Curve of Three Methods with 50 Generations

#### 4. Conclusion

This paper uses SVM to classify the TM remote sensing image of Ningxia Shahu wetland. In order to improve the classification accuracy, it chooses the hybrid particle swarm algorithm to optimize the SVM. In order to make a contrast in performance, the other three methods, namely, the grid search algorithm, genetic algorithm and the particle swarm algorithm are introduced. The comparison and classification results show that GAPSO\_SVM has a better performance than the other three methods in classification accuracy and classifies the remote sensing image more accurately.

In the process of classification and comparison test, some problems that need to be improved. In this paper, spectral feature is applied to do the image classification research, more feature vectors such as texture features should be added to improve the classification performance in the future applications.

#### Acknowledgement

This work is supported by National Natural Science Foundation of China (61162013), the Independent research foundation of Beifang University of nationalities (2011SDJ007), the planned science and technology Project of Ningxia Province under Grant 2010 year, the scientific research projects of State Ethnic Affairs Commission (12BFZ009).

#### References

- [1] ZANG Shu-ying, ZHANG Ce, ZHANG Li-juan, ZHANG Yu-hong. Wetland Remote Sensing Classification Using Support Vector Machine Optimized With Genetic Algorithm. *Scientia Geographica Sinica*. 2012; 32(4): 434-441.
- [2] MelGani F, L Bruzzone. Classification of Hyperspectral Remote Sensing Images with Support Vector Machines. *IEEE Transactions on Geoscience and Remote Sensing*. 2004; 42(8): 1778-1790.

- [3] Dan Liu, Jianming Chen, Guangmin Wu, Haijun Duan. *SVM-Based Remote Sensing Image Classification and Monitoring of Lijiang Chenghai*. *International Conference on Remote Sensing, Environment and Transportation Engineering (RSETE)*. Nanjing: IEEE. 2012:1-4.
- [4] Xianfeng Song, Zheng Duan, Xiaoguang Jiang. Comparison of artificial neural networks and support vector machine classifiers for land cover classification in Northern China using a SPOT-5 HRG image. *International Journal of Remote Sensing*. 2012; 33(10): 3301-3320.
- [5] Yuan Ren, Guangchen Bai. Determination of Optimal SVM Parameters by Using GA/PSO. *Journal Of Computers*. 2010; 5(8): 1160-1168.
- [6] SHAO Xin-guang, YANG Hui-zhong, CHEN GAng. Parameters selection and application of support vector machines based on Particle swarm optimization algorithm. *Control Theory & Applications*. 2006; 23(5): 740-743.
- [7] Muhamad H, Soebagio, A Soeprijanto, MH Purnomo. Modeling and Control of Excitation And Governor Based On Particle Swarm Optimization For Micro Hydro Power Plant. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(2): 297-304.
- [8] Qinghai Bai. Analysis of Particle Swarm Optimization Algorithm. *Computer and Information Science*. 2010; 3(1): 180-184.
- [9] Lovjerg M, Rasmussen TK, Krink T. *Hybrid Particle swarm optimizer with breeding and subpopulation*. IEEE International Conference on Evolutionary Computation. San Diego, USA. 2000:1-6.
- [10] Bing Wang, Xiaoli Wang. Acoustic Performance of Exhaust Muffler Based Genetic Algorithms and Artificial Neural Network. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2013; 11(2): 313-320.
- [11] Cristianini NJ, Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge: Cambridge University Press. 2000.
- [12] LUO Jian\_cheng, ZHOU Cheng\_hu, LEUNG Yee, MA Jiang\_hong. Support Vector Machine for Spatial Feature Extraction and Classification of Remotely Sensed Imagery. *Journal of Remote Sensing*. 2002; 6(1): 50-55.
- [13] HE Ling-min, SHEN Zhang-quan, KONG Fan-sheng, LIU Zhen-ke. Study on Multi-source Remote Sensing Images Classification with SVM. *Journal of Image and Graphics*. 2007; 12(4): 648-654.
- [14] Yuhui Shi, Eberhart R. *A modified particle swarm optimizer*. IEEE International Conference on Evolutionary Computation. Anchorage. Alaska. 1998; 69-73.