

# Hyperspectral Image Classification Based on Novel Binary Particle Swarm with Mutation Mechanism for Band Selection

Lishuan Hu<sup>1,2</sup>, Qun Wang<sup>1,\*</sup> and Tingyan Xing<sup>1</sup>

<sup>1</sup>School of Information Engineering, China University of Geosciences (Beijing), Beijing 100083, China

<sup>2</sup>College of Urban Rail Transit and Logistics, Beijing Union University, Beijing 100101, China

\*Corresponding author

**Abstract**—Hyperspectral remote sensing sensors can capture hundreds of narrow contiguous bands and provide plenty of valuable information. Duo to the high-dimension characteristics of hyperspectral data, band selection plays an important role in the field of Hyperspectral Image (HSI) classification. In this paper, a HSI classification method based on Novel Binary Particle Swarm Optimization with mutation mechanism (MNBPSO) for band selection is proposed. First, we present a thorough experimental study to show the superiority of the MNBPSO method. Then we introduce a pre-processing method for feature selection and parameters determination simultaneously based on MNBPSO. Experiments are conducted on the Indian Pines dataset. The evaluation results show that the proposed approach can select those bands with more discriminative information and improve the classification accuracy effectively.

**Keywords**—binary particle swarm optimization; band selection; parameters determination; mutation mechanism; support vector machines

## I. INTRODUCTION

Hyperspectral Image (HSI) analysis has been an emerging research topic in recent years. Hyperspectral sensors divide the electromagnetic spectrum into hundreds of spectral bands, which can provide the potential and detailed land-cover distinction and identification [1]. The highly correlated spectral bands of hyperspectral images contain a large amount of redundancy, which brings extra time-consuming for hyperspectral data transmission, storage and processing. Additionally, these highly spectral bands may produce the "Hughes phenomenon" [2] and induce overfitting estimation for hyperspectral image classification [3]. Therefore, dimensionality reduction is crucial for hyperspectral image classification. The commonly-used dimensionality reduction methods can be divided into two categories: feature selection [4] and feature extraction [5]. With feature extraction methods, such as principal component analysis (PCA) [6], the high-dimensional feature space is mapped to a low-dimensional space by linear or nonlinear transformations. As a result some crucial information may have been distorted [7]. Otherwise, feature selection methods select appropriate bands from the original set of spectral bands that can preserve the physical meaning of original spectral bands. In this paper, attention is

focused on band selection methods. Feature selection methods apply a set of statistical decision rules, such as the greedy search method [8]. Recently, many heuristic approaches have played an important role in HSI bands selection. Yang et al. [9] proposed a PSO-based dimensionality reduction approach, using the searching criterion function of minimum estimated abundance covariance. Daamouche et al. [10] adopted PSO to select the most informative morphological profiles features for classification. Ghamisi and Benediktsson [11] proposed a new hybridization method for bands selection by integrating both GA and PSO.

In this paper, an improved dimensionality reduction method is presented based on a novel BPSO with mutation mechanism (MNBPSO). The proposed method can select features and determinate Support Vector Machines (SVM) parameters simultaneously effectively. Experiments performed on the Indian Pines dataset demonstrate the effectiveness of the proposed method.

The remainder of this paper is organized as follows. In Section II, we review Particle Swarm Optimization (PSO), Binary Particle Swarm Optimization (BPSO). In Section III, we formulate the framework of our proposed method. We present the experimental results in Section IV, and conclude our work in Section V.

## II. PRELIMINARIES: PSO AND BPSO

### A. PSO

PSO is an evolutionary optimization algorithm originally developed by Kennedy and Eberhart [12], which inspired by observation of social behaviors of animals (e.g., bird flocking). In PSO, a population of particles are used to correspond to individuals. Given that  $p\_best_i$  is the best fitness value recorded so far by the  $i^{th}$  particle in all the preceding iterations. Moreover, the  $g\_best$  is the best position recorded so far by the entire swarm. The velocity and the position of each particle in a d-dimensional space can be updated according to the following formulas:

$$v_{id}(t+1) = w * v_{id} + c_1 r_1 (p\_best_{id} - x_{id}) + c_2 r_2 (g\_best_d - x_{id}) \quad (1)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (2)$$

Where  $w$  is the inertia weight which is constrained to the interval  $[0.0, 1.0]$  and can be adjusted in the direction of linear decrease;  $c_1$  and  $c_2$  are learning rates and  $r_1$  and  $r_2$  are two random functions in the range  $[0, 1]$ . The termination criterion for iterations is determined according to whether the maximum generation or a designated value of the fitness is reached.

### B. BPSO

PSO is operated in continuous space. But many optimization problems are set in discrete space. So Kennedy and Eberhart extended the PSO to a binary space (BPSO) in 1997 [13]. In BPSO, a particle moves in a state space restricted to 0 and 1 on each dimension, where  $v_{id}$  represents the probability of the bit  $x_{id}$  taking the value 1. As Figure 1 show, a logistic transformation ( $v_{id}$ ) can be used to accomplish this modification, and the position will be updated according to equation (3) and (4).

$$S(v_{id}) = \frac{1}{1 + e^{-v_{id}}} \quad (3)$$

$$x_{id} = \begin{cases} 1 & \text{rand}() < S(v_{id}) \\ 0 & \text{else} \end{cases} \quad (4)$$

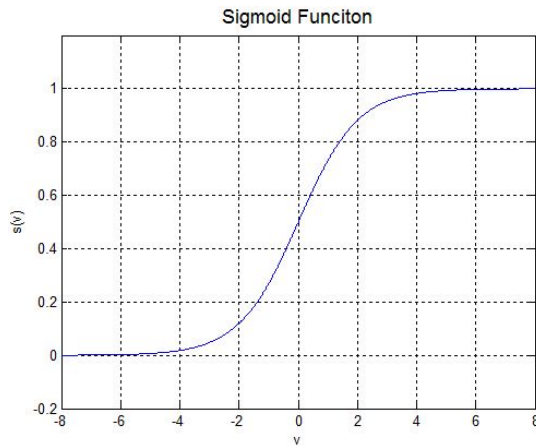


FIGURE I. BPSO SIGMOID FUNCTION

### III. PROPOSED METHOD

In this section, we address the proposed method: a novel BPSO with mutation mechanism (MNBPSO) for HSI classification bands selection and parameters determination simultaneously.

#### A. A Novel BPSO with Mutation Mechanism

In the standard PSO, a big value of  $v_{id}$  in the positive and negative direction means the particle requires a greater

movement based on its previous position. But in BPSO, the bigger value in the positive and negative direction causes bigger probability of 1 and 0 for the particle position respectively. Also, in the standard PSO,  $v_{id}$  goes to zero means the particle has a suitable position and does not move again. While, in BPSO, according to sigmoid function,  $v_{id}$  goes to zero means the position may be changed by taking the value of 1 or 0 with the probability of 0.5. This causes the algorithm not to converge well.

To overcome these disadvantage associated with the BPSO, a novel BPSO (NBPSO) is presented by Nezamabadi-pour and Rostami-shahrbabaki [14]. In the algorithm a proper probability function can be defined as equation (5). The position will be updated according to equation (6).

$$S'(v_{id}) = 2 \times \left| \frac{1}{1 + e^{-v_{id}}} - 0.5 \right| \quad (5)$$

$$x_{id}(t+1) = \begin{cases} \text{exchange}(x_{id}(t)) & \text{rand}() < S'(v_{id}) \\ x_{id}(t) & \text{else} \end{cases} \quad (6)$$

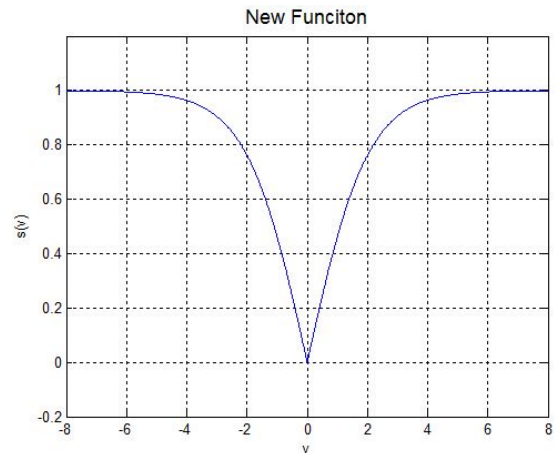


FIGURE II. NBPSO NEW FUNCTION

Figure 2 presents a depiction of NBPSO function. A big value for  $v_{id}$  means the position is not good and it changes from 0 to 1 or vice versa with a high probability. Also a small value for  $v_{id}$  shows it changes the position with a low probability and when  $v_{id}$  becomes zero, the position will remain unchanged.

The same as standard PSO, NBPSO may converge at the early stage, it usually sinks into the local optimum during the later stages of the optimization. This phenomenon is referred to as stagnation.

To overcome the premature convergence and slow convergent speed, the mutation mechanism is adopted to the NBPSO (MNBPSO) to enhance the global search capability. In this way, when the global optimal solution does not improve

with the increasing number of generations,  $p$  percent of particles and  $q_i$  percent dimensions of the  $t^{\text{th}}$  particle are selected randomly to do mutation operation.

```
for (i=1; i<=N*p; i++){
    t=(int)(rand()*N);
    for (j=1; j<=D; j++){
        if (rand()<qi) then
            if (x(t, j)==0) then x(t, j)=1;
            else x(t, j)=0;
        }
    }
}
```

Seven benchmark [15] functions are use to verify the MNBPSO method. Figure 3 and Figure 4 show the convergence characteristics for one of these benchmark functions. Other functions can get the similar results. These results show that MNBPSO can achieve the best result.

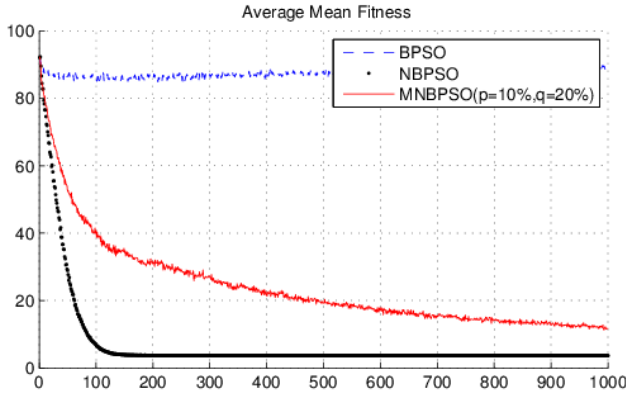


FIGURE III. CONVERGENCE CHARACTERISTICS: AVERAGE MEAN FITNESS

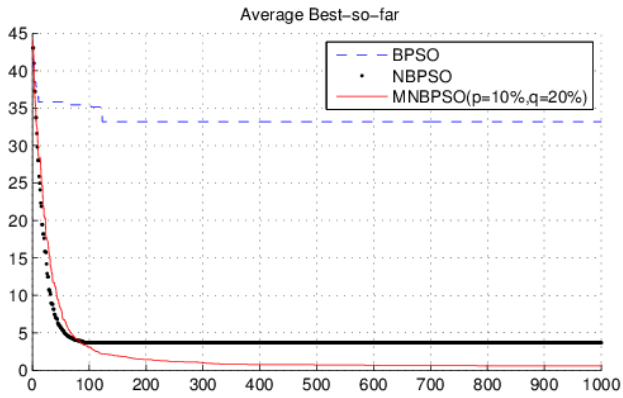


FIGURE IV. CONVERGENCE CHARACTERISTICS: BEST-SO-FAR

#### B. Feature Selection and Parameters Determinationnits Based on MNBPSO

The nonlinear SVM classifier based on the Gaussian kernel is proved has a promising generalization performance for supporting the HSI classification [16]. But, it needs to choose

proper kernel parameters  $c$  and  $\gamma$ . Because of the good performances generally achieved with our improved MNBPSO, we propose an algorithm based on MNBPSO for parameters determination and feature selection simultaneously.

In the proposed approach, each particle consists of two part bits. One part is for feature selection, which has  $D$  bits according to  $D$  features, the value of each bit is either 1 or 0, which respectively indicates whether the corresponding feature is selected or not. The other part is for parameter determination, in which there are  $L$  bits for each parameter ( $c$  and  $\gamma$ ). All of the  $L$  bits are 0 indicates the minimums values of the parameter, vise versa, all bits are 1 indicates the max value of the parameter. In this paper  $L$  is set to be 15. A transform function is developed to obtain the values of these parameters from inputting bits string.

As shown in Figure 5, the total bits of a particle is  $D+2L$ .

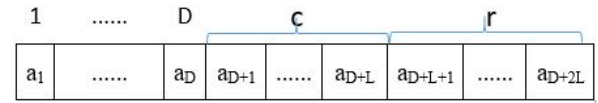


FIGURE V. SOLUTION REPRESENTATION

The fitness function in this study is the classification accuracy. Finally the particle's velocity and position are updated based on the proposed method MNBPSO until the termination condition is satisfied.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSION

##### A. Data Set Description

The India Pines dataset was collected by the AVIRIS sensor over the Indian Pines region, Northwestern Indian, USA, in 1992, which is widely used to verify the performance of classification algorithms. The Indian Pines data set comprises 220 bands with the spatial size of 145\*145 pixels. Removing the noisy bands, 200 bands remained. The ground truth has 10,062 labeled pixels which consists of 16 land cover classes. Figure 7(a) shows the ground truth.

In our experiment, small sample size classes are removed, only 9 classes are considered. Table 1 summarizes these considered classes and the number of training and testing samples for these classes.

##### B. Experimental Result and Discussion

According to our proposed method, bands selection and SVM parameters determination is performed simultaneously. We compare MNBPSO with BPSO to verify our proposed method. Figure 6(a), (b) and (c) show the selected bands preserve the same style as the original bands. Figure 6(d) shows MNBPSO can avoid sinking into the local optimum during the later stages and get the better result than BPSO.

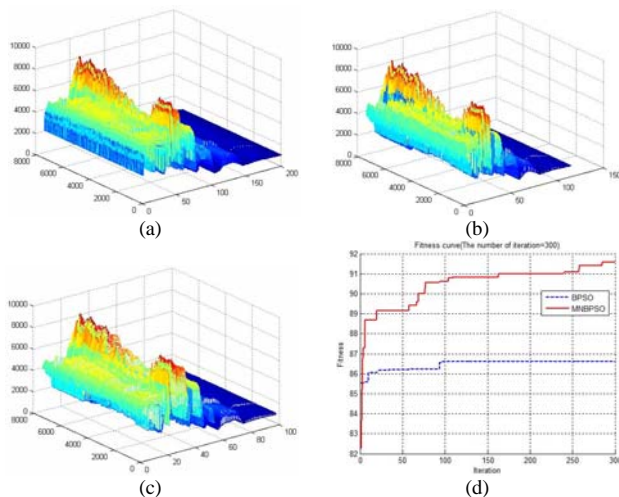


FIGURE VI. BANDS SELECTION OF THE AVIRIS INDIAN PINE DATASET: (A) THE TOTAL BANDS, (B) BANDS SELECTED BY BPSO, (C) BANDS SELECTED BY MNBP-SO, (D) EVOLUTIONARY PROCESS OF BPSO AND MNBP-SO.

Three evaluation metrics, overall accuracy (OA), average accuracies (AA), and Kappa coefficient, are widely used to measure the statistical significance for hyperspectral image classification. From Figure 7 and table 1, we can easily see that, our proposed method, SVM+MNBP-SO, obtains a higher accuracy compared with SVM and SVM+BPSO. It is clear that SVM+MNBP-SO can obtain the better performance.

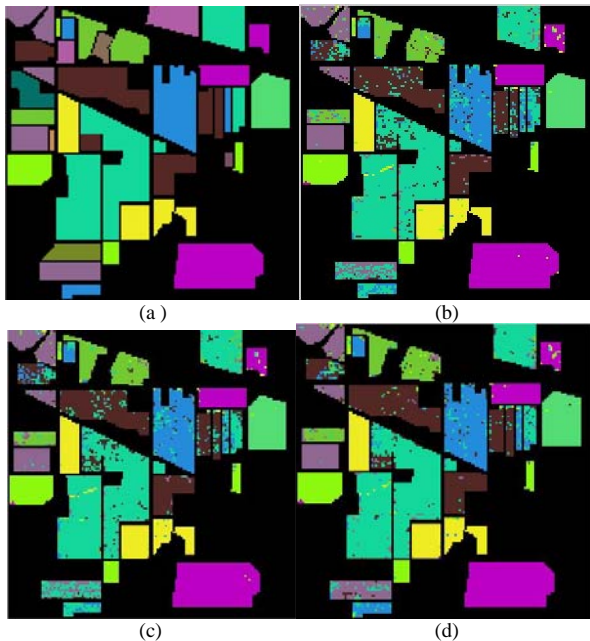


FIGURE VII. CLASSIFICATION MAPS OF THE AVIRIS INDIAN PINE DATASET: (A) GROUND TRUTH, (B) CLASSIFICATION MAP FOR THE SVM METHOD, (C) CLASSIFICATION MAP FOR THE SVM+BPSO, (D) CLASSIFICATION MAP FOR THE SVM+MNBP-SO.

TABLE I. GLOBAL AND CLASS-SPECIFIC CLASSIFICATION ACCURACIES IN PERCENTAGE

Indian Pines	Training /Test	SVM	SVM +BPSO	SVM +MNBP-SO
Corn-notill	304/1124	79.09	78.91	<b>88.61</b>
Corn-mintill	157/637	73.85	71.47	<b>87.07</b>
Grass/Pasture	91/392	95.66	95.92	<b>96.94</b>
Grass/Trees	137/593	96.80	97.64	<b>98.48</b>
Hay-windrowed	91/387	99.48	99.22	<b>99.71</b>
Soybeans-notill	187/85	70.45	71.97	<b>85.73</b>
Soybeans-mintill	511/1944	85.80	87.71	<b>90.28</b>
Soybeans-clean	106/487	80.49	<b>88.50</b>	86.86
Woods	235/1030	97.96	<b>98.45</b>	98.16
selected bands	-	200	110	<b>97</b>
Kappa	-	82.95	84.20	<b>90.12</b>
OA	-	85.53	86.61	<b>91.57</b>
AA	-	86.62	87.76	<b>92.32</b>

## V. CONCLUSION

In this paper, a heuristic approach about MNBP-SO has been applied to hyperspectral remote sensing image classification. Our proposed method can select the bands and determinate SVM parameters simultaneously effectively. Experimental results show that the proposed method obtain more accurate results than the SVM, SVM+BPSO. The algorithm is a promising method on HSI classification.

## ACKNOWLEDGEMENT

The authors would like to thank D.Landgrebe and L.Biehl for making the AVIRIS Indian Pines data available to the community and C.C. Chang and C.J. Lin for making SVM toolkit openly available.

## REFERENCES

- [1] C.I. Chang, Hyperspectral Data Exploitation: Theory and Applications, Hoboken, NJ, Wiley, USA, 2007.
- [2] G.F.Hughes, On the mean accuracy of statistical pattern recognizers, IEEE Trans. Inf. Theory, vol. 14, no. 1, pp. 55-63, Jan. 1968.
- [3] G.Licciardi, F.D.Frate, and R.Duca, Feature reduction of hyperspectral data using autoassociative neural networks algorithms, in Proc.IEEE Int.Geosci.Remote Sens. Symp., Jul. 12-17,2009, vol. 1, pp. 176-179.
- [4] Xue, B., Nguyen, S., Zhang, M.: A new binary particle swarm optimisation algorithm for feature selection. In: European Conference on the Applications of Evolutionary Computation. pp.501-513. 2014.
- [5] A.Bal, Image feature extraction by dynamic neural filtering and phase-only joint transform correlation, Opt. Laser Technol. 39(1) . 2007).
- [6] S.Wold, K.Esbensen, P.Geladi, Principal component analysis, Chemometr.Intell.Lab.2(1-3).pp. 37-52.1987.
- [7] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989..
- [8] L.Yu, H.Liu, Efficient feature selection via analysis of relevance and redundancy, J.Mach.Learn.Res.5.pp.1205-1224 2004.
- [9] H. Yang , Q. Du , G. Chen , Particle swarm optimization-based hyperspectral di- dimensionality reduction for urban land cover

- classification, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens* 5, pp. 544–554. 2012 .
- [10] A. Daamouche , F. Melgani , N. Alajlan , N. Conci , Swarm optimization of structuring elements for VHR image classification, *IEEE Geosci. Remote Sens. Lett.* 10, pp. 1334–1338 . 2013.
  - [11] P. Ghamisi , J.A. Benediktsson , Feature selection based on hybridization of genetic algorithm and particle swarm optimization, *IEEE Geosci. Remote Sens. Lett.* 12, pp. 309–313. 2015.
  - [12] Kennedy J, Eberhart R C, Particle swarm optimization, In: *Proceedings of 1995 IEEE International Conference on Neural Networks*, Perth, Australia, pp. 1942–1948. 1995.
  - [13] Kennedy J, Eberhart R C, A discrete binary version for the particle swarm optimization, In: *Proceedings of 1997 IEEE International Conference on System, Man, and Cybernetics. Computational Cybernetics and Simulation*, Orlando, Florida, USA, 1997.
  - [14] H. Nezamabadi-pour, M. Rostami-shahrabaki, M.M. Farsangi, “Binary Particle Swarm Optimization: challenges and New Solutions”, *The Journal of Computer Society of Iran (CSI) On Computer Science and Engineering (JCSE)*, vol. 6, no. (1-A), pp. 21-32, 2008.
  - [15] Digalakis J.G, and Margaritis K.G. An experimental study of benchmarking functions for genetic algorithms. *International Journal of Computer Mathematics*, 79(4), pp.403-416. 2002.
  - [16] E. Pasolli, F. Melgani, D. Tuia, F. Pacifici, W.J. Emery, SVM active learning approach for image classification using spatial information, *IEEE Trans. Geosci. Remote Sens.* 52 .pp. 2217-2233. 2014.