Hyperspectral Dimensionality Reduction of Forest Types Based on Cat Swarm Algorithm

Li Yan[1], XingYan-Qiu[2] and Wang Li-Hai[2,*]

[1] *Institute of Information and Computer Engineering, Northeast Forestry University, Harbin, 150040, P. R. China*

[2] *Forest Operations and Environmental Research Center Northeast Forestry University, Harbin 150040, P. R. China*

Abstract: One of the main ways of dimensionality reduction of hyperspectral image was band selection. The paper proposed a hyperspectral image bands selection method based on binary cat swarm algorithm to solve problems of the high complexity and intensive computation efficiently for follow-up applied research. In this paper, Jilin Wangqing Forestry Bureau was chosen as the study area, by optimization process of the cats' location electing, less associated and more informative bands were selected from 115 bands of HJ-1A, band combination (22,37,109), to distinguish 5 kinds of dominant tree species and get better classification accuracy.

1. INTRODUCTION

The hyperspectral remote-sensing technique has been widely used in various fields due to its capacity to handle large data amount, high spectral resolution, wide range, narrow band, massive information, *etc*. [1], but its further application is restricted by its "dimensionality curse" brought by massive data [2] and processing difficulty. Prior to data application, it is necessary to reduce the dimensionality in both mainstream ways: band selection and feature extraction ways [3]; and the most effective feature or band beneficial to the subsequent classification is found from the highly-dimensional feature space. Numerous researchers researched on the hyperspectral remote-sensing dimensionality reduction partially, but their used data or methodologies were limited. Zhou Shuang, *et al*. reduced the dimensionality for AVIRIS data using the ant colony algorithm [4], however they researched on the crops in American Indian areas and the ant colony algorithm is prone to fall within the locally optimal solution. Wang Jiwei *et al*. [5] combined the bands of HJ-1A data using the method of Optimum Index Factor (OIF). However, they were mainly based on the band information amount, and didn't select the band in the perspective of spectral feature of ground objects. Li Junming *et al*. [6] studied the HJ-lA hyperspectral imaging dimensionality reduction by comparing two methods: OIF and Optimum bands combination; however these methods are of much calculation amount and lack of intelligence.

The present paper focuses on the band selection methodology, looks for and combines the bands with large ground object spectrum difference, and less band correlation and large information amount. The present paper uses the bionic cat swarm algorithm, an emerging swarm intelligent algorithm, establishes a structure suitable for the hyperspectral imaging and finds an excellent band combination, eliminates the data with high correlation to the band and serious data redundancy, reduces the subsequent classification processing time, improves the classification accuracy and efficiency [7, 8], and distracts the manual intervention in band selection.

1.1. Study Area

The study area in the present paper is the operating area of Wangqing Forest Bureau, southeastern Wangqing County, northeastern Yanbian City, Jilin Province. The area is located at 129°56'~ 131°04'E and 43°05' ~ 43°40'N covering a total area of 0.304 million hm^2 and a wooded area of 0.187 million hm^2. It is a typical natural secondary forest with abundant resource that contains the mixed zonal- distributed broadleaf-conifer trees. The conifers mainly include Pinus koraiensis, Picea, Symplocarpus Salisb, Larix kaempferi, broadleaf include Fraxinus mandschurica, Juglans mandshurica, Quercus mongolica, Tilia, Acer mono, Ulmus pumila, Betula platyphylla, Populus and Betuladavuric, *etc*. [9]. And the main pure forest tree species are Mongolian oak, birch and larch that make up for over 90% of the trees.

1.2. Basic Data

1.2.1. Hyperspectral Data

The Hyperspectral Imaging Radiometer (HSI) is carried on HJ-1A/B, with an image width of over 50 kms. The ground pixel resolution is 100 m and a total of 115 bands are set, which include blue bands (1-6), cyan bands (7-19),

*Address correspondence to this author at the Northeast Forestry University, He Xing Road, Harbin, 150040, P. R. China;
E-mail: woshimao2011@163.com

green bands (20-40), yellow bands (41-49), orange bands (50-58), red bands (59-88) and near-infrared bands (89-115). The spectral range is 0.45-0.9 um and the average spectral resolution is 4.32 nm. The $\pm30^{\circ}$ side-viewing ability and on-orbit calibration function are provided. The revisit period is 96 hrs [10]. The hyperspectral data processed in the present paper are the images snapped on June 9, 2010 and the spectral image SN is 321381. The data is a level-2 data, of which 1~20 and 113, 114 and 115 bands are bad data, so the 92 bands are actually analyzed; 21~88 bands are the visible part and 90~112 are near-infrared part. The data is corrected in terms of logarithmic residue atmospherically, and then corrected the accuracy for the ground control points geometrically to form the basic data for further best band combination research.

1.2.2. Field Survey Data

In the present study, the field surveys were conducted respectively in September 2007 and September 2010. 86 round plots with horizontal projected area of 500 m² were randomly set up according to the forest types, slopes, *etc*. The standing trees within the plots were measured, each including DBH, tree height, tree species and canopy density. The trees for field survey are divided into 5 groups: Mongolia oak, birch, spruce, poplar, and larch. Every dominant tree species accounted for over 65% in the accumulation and the numbers of plots were 30,19,15,12 and 10, respectively [11].

1.3. Basic Thought of Cat Swarm Optimization

The Cat Swarm Optimization (CSO) algorithm is an intelligent optimization algorithm that is built up to simulate the behavior model of individual cats and cat swarm in daily life in recent years. The algorithm is still at the study phase. The present paper is intended to initially explore the application of the excellent method in the remote-sensing field. The cat location information in the algorithm is the possible solution of the problem to be optimized. In the algorithm, the main activities of cats are divided into search and tracing as being respectively appropriate to both dissimilar patterns in the algorithm [12].

1.3.1. Seeking Mode

In this mode, every cat's self-location copy is applied with a deviation operator to produce a new location (neighborhood), the copycat is placed in the memory pool, and the fitness is calculated [13]. In the memory pool, a candidate point of the highest fitness is selected as the next location point to which the cat will move, thus achieving the location update. The current cat location is updated with the global optimal location. During an iterative process, the calculation is done according to the fitness function and the current optimal solution is reserved. Parameters under the mode include SMP, SRD and CDC. SMP: memory pool defines the scope observable to each cat, also the searchable memory size available to each cat; and storing the searchable location points to a cat. The cat will select a best location point from the memory pool according to the fitness magnitude; SRD: change domain of selected dimension; CDC: change ratio of dimension.

1.3.2. Tracking Mode

In this mode, the cat moves at self-speed. Each gene value is moved using the speed - displacement model [14]. The best location underwent by the whole cat swarm is the optimal solution $x_{best,d}(t)$ that is searchable at time *t*, namely, the current cat speed is updated by the global optimal location, and then the current cat location is changed by the updated speed value to continuously direct to the best optimal solution [15]. The process is as follows:

(1) From the formula (1), to change the speed $v_{k,d}$ of each dimension.

$$v_{k,d}(t+1) = v_{k,d}(t) + c \times rand \times (x_{best,d}(t) - x_{k,d}(t)) \qquad d = 1,2,3\cdots,D \qquad (1)$$

$v_{k,d}(t)$ is the speed of the kth cat at the dth gene bit. Each cat speed is $V_i = \{V_i^1, V_i^2, \ldots, V_i^L\}$, *L* is the total gene length of the ith cat. *c* is a constant; and *r* is random number between 0~1. The value of current cat's gene bit at the same location is replaced with the dth gene bit in the cat with the highest fitness value so as to update the current cat speed.

(2) Based on the previous step, to update the current cat using the equation (2), i.e. the kth cat's location.

$$x_{k,d}(t+1) = x_{k,d}(t) + v_{k,d}(t+1) \qquad (2)$$

2. PROPOSAL OF BAND EXTRACTION USING CATS SWARM ALGORITHM

2.1. Structural Design

Each band of hyperspectral image can be viewed as a feature carrying some information. The band selection is actually a feature selection in the high-dimensional feature space. The band of hyperspectral image has more sub-sets, and has more band combination modes. Selecting a band combination with large amount of information, less correlation and strong class separability is a combined optimization problems. In the present paper, Binary Cat Swarm Optimization (BCSO) is proposed to solve the discrete problem. Searching a feature band of hyperspectral image is a process of searching with the algorithm.

The main thought of this algorithm is that the cat location value of each gene bit is defined as 1-dimensional non-zero value, i.e. 1. With automatic search, it draws an optimal solution of cat swarm, and a band combination to be searched. For each gene bit in corresponding individual cat, 0 indicates that the band is not selected and 1 indicates that it is selected.

Each cat has total D-dimensional data. The D-bit binary data indicates the band mask. The encoding structure for each cat is illustrated as in Fig. (**1**).

2.2. Fitness Selection

Hyperspectral band selection is primarily on the basis of the amount of information carried by the bands, inter-band correlation and class separability, namely, the characteristic

Gene bit 1 Gene bit 2 .... Gene bit D-1 Gene bit D

Location code	$x_{i,1}$	$x_{i,2}$	......	$x_{i,n-1}$	$x_{i,n}$
Speed code	$v_{i,1}$	$v_{i,2}$	......	$v_{i,n-1}$	$v_{i,n}$

Fig. (1). Representation of band subset based on BCSO encoding. Each cat's location is $X_{k,d}(t)$, k=1,2…,i, d=1,2…,n. The algorithm gives dissimilar band combinations iteratively several times. The targeted function's optimal solution is then given by combining with the classification accuracy evaluation.

of spectral curve in ground object is made as a criterion [16]. The single criterion can be used for the fitness calculation, however, the present paper tried to calculate the fitness by combining different criteria. The fitness level decides the probability of individuals being selected. The present paper combined the joint entropy of combined band images and Jeffries-Matusita Distance as the fitness equation and as criteria for the amount of information and class separability in the band combination. This may reduce the sidedness owing to the band selection by only one criterion.

(1) Joint entropy

According to information theory, the image f (x, y) entropy is:

$$H(f(x,y)) = -\sum_{i=0}^{255} P_i \log_2 {}^{P_i} \qquad (3)$$

P_i is the probability of the image at gray pixel value i, which is derived from the image histogram. The joint entropy characterizes the amount of image's information in combined band.

(2) J-M distance

J-M distance characterizes the separability between two classes. For any given ground object classes, it is based on the spectral separability indicator in the conditional probability theory. The equation is as follows:

$$J_{ij} = \left\{ \int_x \left[\sqrt{p(\frac{X}{\omega_i})} - \sqrt{p(\frac{X}{\omega_j})} \right]^2 dX \right\}^{\frac{1}{2}} \qquad (4)$$

$p(\frac{X}{\omega_i})$ is the conditional probability density, i.e. the probability of ith pixel belonging to the ω_ith class. J_{ij} value is between 0~2 and its magnitude represents the separability extent among samples. When J_{ij} is located between 1.9~2.0, there is a good spectral separability among samples.

(3)Fitness equation

After the individuals obtained from both indicators, joint entropy and J-M distance, are normalized, the fitness equation is created as below:

$$fitness = \alpha \cdot H^* + \beta \cdot J^* \qquad (5)$$

H^* and J^* are dimensionless values after the joint entropies and J-M distance are normalized; α and β are the values inside [0,1].

2.3. Implementation of Band Selection Based on Modified Discrete Cat Swarm Algorithm

The BCSO-based band selection algorithm is as follows:

Step 1: Set parameters. To set the maximum iteration number MaxIter, SMP, SRD, CDC, and grouping rate MR. The cat swarm is divided into seeking mode or tracing mode according to MR.

Step 2: Initialize the cat swarm. To set the cat swarm size m, each cat dimension D, D=n. n is the number of hyperspectral bands. Each dimensional location $x_{i,d}$, either 0 or 1, of the ith cat is randomly initialized. The speed $v_{i,d}$ is randomly initialized for each dimensional location. If the hyper spectrum is divided into s sub-sets according to the correlation coefficient matrix in advance, then the location value of a gene bit in each band sub-set is 1 by being set randomly in individual cat and the rest gene bit is 0, namely, in the cat's D gene bits, there should be randomly s location values as 1($s \leq D$).

Step 3: to calculate the fitness for each cat according to Equations (3), (4) and (5).

Step 4: the cat swarm is divided into seeking and tracing modes according to the grouping rate.

Step 5: To set a flag bit for each cat, determining which pattern the cat belongs to according to the flag bit. If at seeking mode, go to Step 6, or otherwise, execute the tracing mode and go to Step 7.

Step 6: go to seeking mode.

Step 6.1: to view the random band combination as a cat, copy itself by SMP and put them in the memory pool.

Step 6.2: to update the location according to the mutation rate CDC and change domain SDR for each cat, determine which dimensional locations requires changes for the cat by CDC value, and also set a gene bit that does not involve in mutation. The mutation for each gene bit is updated as per equation (6).

$$x_{i,d}(t+1) = \begin{cases} 0 & x_{i,d}(t)=1 \\ 1 & else \end{cases} \qquad (6)$$

In BCSO algorithm proposed in the present paper, in line with the cat's encoding structure and problem-solving needs, if the hyperspectral band is divided into s sub-sets according to the correlation coefficient matrix, then the gene bit 1 in each sub-set within the every cat's s sub-sets must be mutated from 1 to 0; and the gene bit 0 is mutated into 1, namely
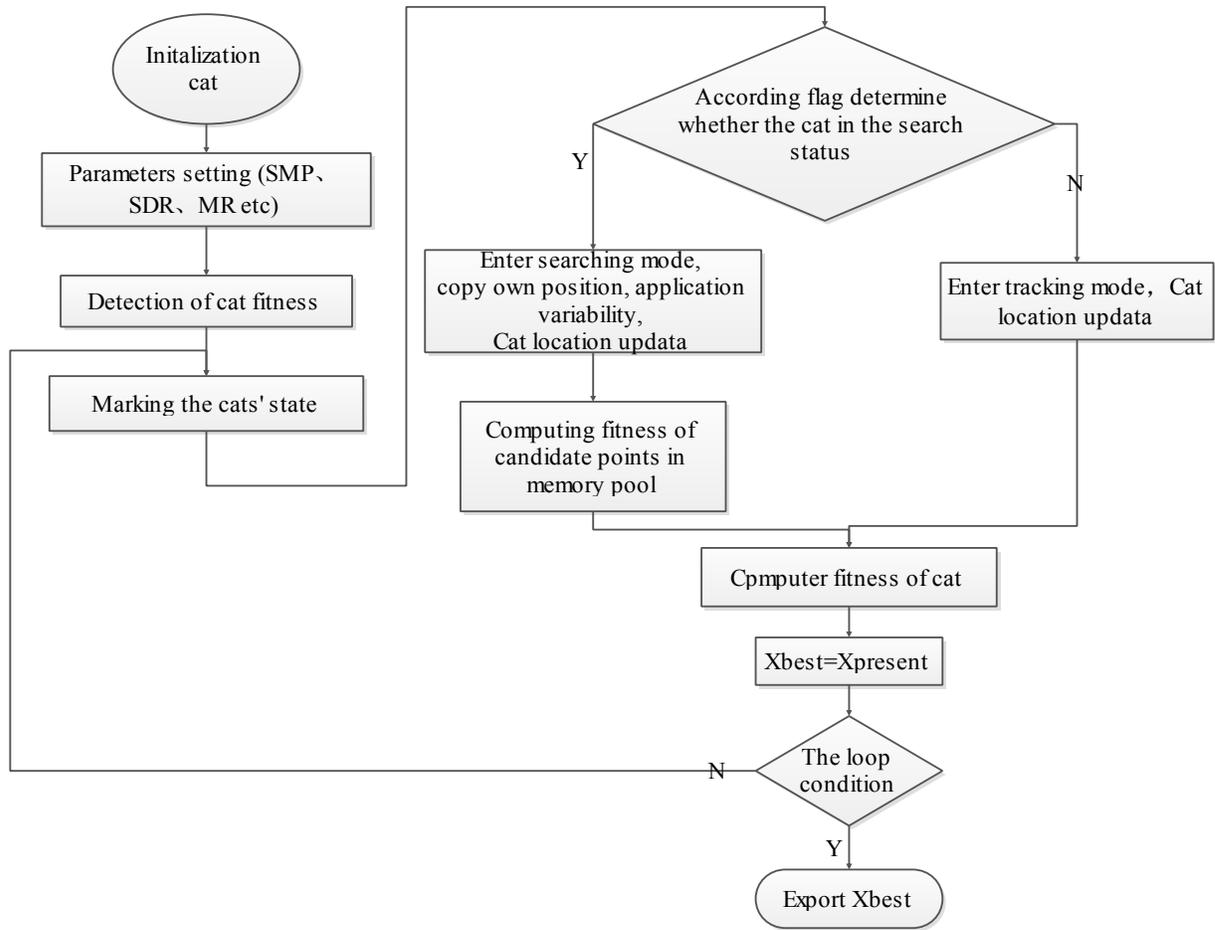
Fig. (2). BCSO-based hyperspectral band selection flowchart.

the mutated gene bit is 2s, when CDC is 1. This will ensure that there is still a bit kept as 1 in each band sub-set after mutation, namely 1 band is selected within the sub-set. Providing CDC is 0.8, the mutated dimensions of individual cat should be 2 * [0.8 * s + 0.5].

Step 6.3: To calculate the fitness of all individuals in the memory pool according to equation (5).

Step 6.4: To select the candidate cat with maximal fitness in place of current cat's location from the copied SMP cats by sorting or selecting the operator.

Step 7: To enter the tracing pattern, equivalent to search globally.

Step 7.1: To calculate fitness, record the global optimal location x_{best} of the cat swarm.

Step 7.2: To update the speed for each cat according to equation (1), update its location according to the formulas (7) and (8), find the globally optimal individuals and reach the optimal location. If the hyperspectral band is divided into several sub-sets, then perform similarly as Step 6.2 to suppress the location update. With continuous iteration, the cats are always approaching gradually towards the optimal solution, namely approaching the optimal band combination.

$$Sig(v_{i,d}) = \frac{1}{1 + \exp(-v_{i,d})} \tag{7}$$

$$x_{i,d}(t+1) = \begin{cases} 0, & r \geq Sig(v_{i,d}(t+1)) \\ 1, & else \end{cases} \tag{8}$$

Where, *Sig* is the *Sigmoid* function that the continuous speed values are discrete as 0 or 1. *r* is a random quantity subject to the uniform distribution. $x_{i,d}(t+1)$ represents the dth component of the ith cat at time *t*.

Step 8: To calculate the fitness for all cats, record and retain the optimal cat after the fitness is converged, namely, the global optimal solution x_{best}. If the suboptimal search is done, record and retain the cats by the former *q* fitness.

Step 9: to determine whether the termination condition is met. If the set maximal iteration is reached, or the solution meets a threshold, quit, or otherwise, continuously execute Step 4.

The band selection flowchart is illustrated as in Fig. (**2**). based on the BCSO algorithm.
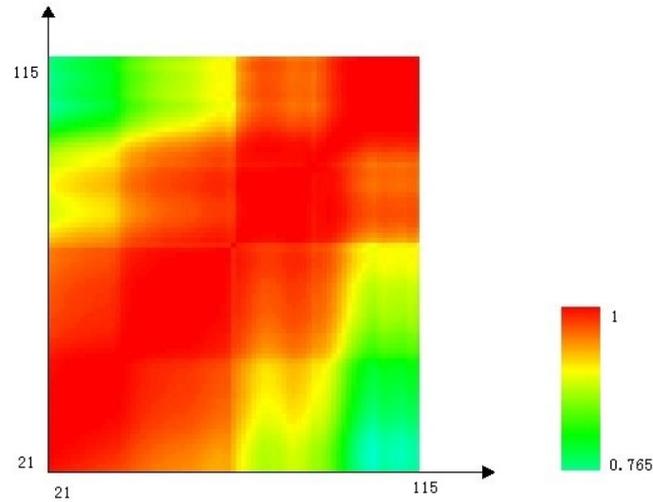
Fig. (3). Band correlation coefficient matrix.

Table 1. Parameter of band selection based on BCSO.

Parameter	Value
m	10
n	92
q	20
MaxIter	20
MR	0.02
SMP	5
SRD	1
CDC	0.8
α	0.7
β	0.3

3. RESULTS AND ANALYSIS

In this paper, 92 noise-removed bands of HSI hyperspectral images from HJ-1A satellite were processed using MATLAB simulation programming. The mean, variance and other basic statistic data are calculated to give the correlation coefficient matrix among bands of images as shown in Fig. (**3**). The brighter colors in the figure indicate that the adjacent image bands are strongly correlated and the correlation coefficient is larger. With the correlation coefficient vectors of adjacent bands derived from the symmetric matrix, the sub-sets were divided and 92 bands were divided into 3 sub-sets of 21~30, 31~76 and 77~112 and the band sub-sets were suppressed using the BCSO algorithm proposed here. 3 bands were selected for combination by searching 3 sub-sets to look for the optimal band combination so that 5 dissimilar tree species, birch, Mongolian oak, poplar, larch and spruce can be distinguished most easily.

Depending on the experiences and problem complexity as well as BCSO algorithm, all parameters for band selection are as shown in Table **1**.

After iterative calculation, top 20 band combinations with higher fitness and optimal solution were derived as shown in Table **2**.

5 tree species were processed based on the spectral characteristics of different ground objects, respectively. To compare the results during experiment, J-M distance was firstly selected as fitness for 20 iterations independently and the iterative result tends to be 1.9037. When the joint entropy makes the fitness for 20 iterations independently, it is not strongly convergent. In the present paper, the fitness after iteration by fitness proposed in the present paper tends to be 1.6472, the cat is having a tendency to move toward the optimal solution, its band combination is (22, 37, 109), and

Table 2. The best band combination.

Band 1	Band 2	Band 3	Fitness
22	37	109	1.6472
22	36	109	1.6552
21	36	109	1.7962
22	32	108	1.7854
22	33	107	1.7804
21	37	109	1.7878
21	32	109	1.9273
21	33	108	1.9220
22	43	109	1.4449
22	38	109	1.6507
22	41	109	1.5024
21	42	109	1.4487
21	43	109	1.5836
22	40	109	1.5440
21	38	109	1.7909
21	41	108	1.6420
22	44	109	1.3672
21	42	108	1.5875
22	35	109	1.6761
21	40	109	1.6841
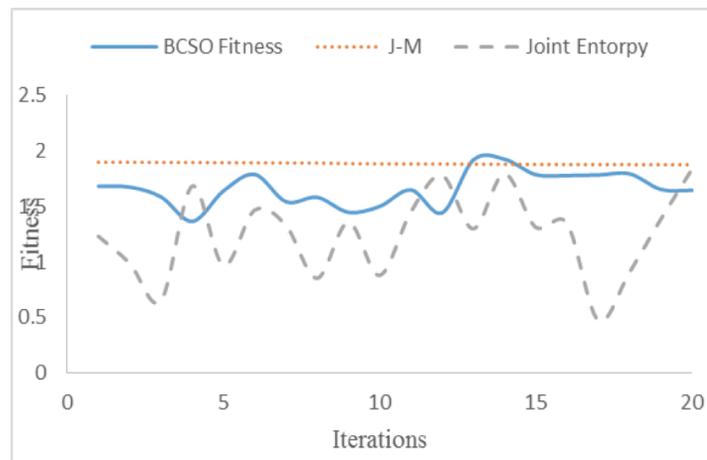
Fig. (4). The relationship comparison chart of five species between iterations and fitness.

such band combination has the best separability for 5 tree species. See Fig. (**4**) for the relation between convergent trend and fitness among fitness of the paper, J-M distance and joint entropy.

When α and β parameters of BCSO fitness proposed in the present paper are changed, the fitness gradually tends to be 1.5703 and the optimal solution is the band combination (22, 43, 110) given α =0.3 and β =0.7, as Fig. (**5**).
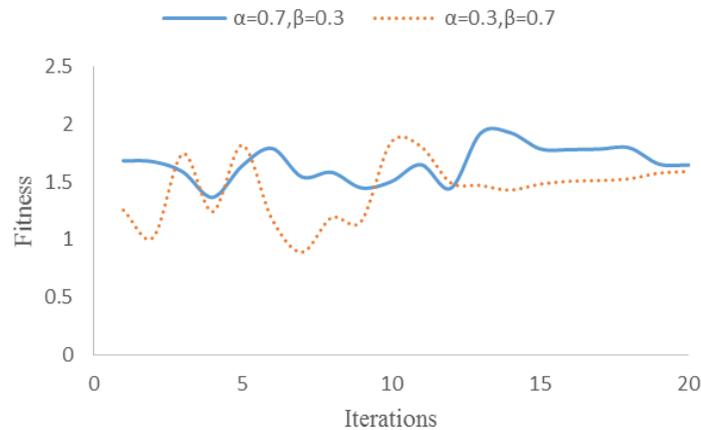
Fig. (5). The relationship comparison chart of five species between iterations and fitness based on different parameters.

Table 3.   Comparison of classification verification results.

Band Combination	Method	Parameter	Classification Precision	Kappa
22,37,109	BCSO proposed here	α =0.7, β =0.3	82%	0.775
22,36,109	Ditto	Ditto	81%	0.7625
22,38,109	Ditto	Ditto	75%	0.6875
21,41,108	Ditto	Ditto	73%	0.6625
22,35,109	Ditto	Ditto	76%	0.7
21,40,109	Ditto	Ditto	72%	0.65
22,43,110	Ditto	α =0.3, β =0.7	78%	0.725
21,43,109	PSO	α =0.7, β =0.3	79%	0.7375

In order to evaluate the optimal band combination, the 5 band combinations approximate to the fitness 1.6472 appropriate to the optimal solution are given in Table **2**, optimal band combination (22,43,110) derived after parameters α and β are changed and the optimal band combination (21,43,109) after band search with identical parameters using the PSO algorithm were selected. The synthetic images were classified and verified, respectively. The classification accuracy and Kappa coefficient are given to analyze the classification results using the confusion matrix. For the classification verification, the plots in the study area were collected as training sets, and the images after band combination were classified using the maximum likelihood (MLC). See Table **3** for the results.

For the band combination (22,37,109), the total classification accuracy is 82%, Kappa coefficient is 0.775; for the band combination (22,43,110), the classification accuracy is 78%, Kappa coefficient is 0.725 which is slightly lower than that of band combination (22,37,109). This complies with the fitness higher than (22,43,110) when the band combination (22,37,109) converges as shown in Fig. (**4**). For the band combination (21, 43, 109), the total classification accuracy is 79 %, Kappa coefficient is 0.7375, suggesting that the optimal solution derived using BCSO algorithm is better than PSO algorithm in the afterward classification. It is concluded from Table **3** that the optimal band combination (22,37,109) derived from BCSO optimization shows the best distinguishing degree for 5 dominant tree species. In comparison to the spectral images of ground objects, it is found that the reflectance values of each species vary significantly at band 22, less at 109 poplar and spruce spectral curves, and largely from other species. This complies with the optimization results of cat swarm algorithm.

CONCLUSION

The present paper improved the cat swarm optimization algorithm, a new swarm intelligence algorithm, based on the basic thoughts, brought forth a discrete BCSO, designed a structure and encoding rule that are applied in the hyperspectral image processing, and finally used them in the remote sensing forestry. The present paper mapped and studied the dimensionality reduction process of hyperspectral data into

an optimization process of BCSO in a band selection way. With spectral characteristics of different ground objects, the band combinations that have weak correlation and are better to distinguish 5 dominant species were selected from 115 bands of HJ-1A remote sensing images. The cat swarm optimization is stronger in self-adaption, self-organization and self-learning and may explore a new path to solve the remote-sensing problems.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Q. Tong, B. Zhang, and L. Zheng, "Hyperspectral Remote-Sensing-Principle, Technology and Application". *Higher Education Press*: Beijing, 2006.

[2]   G. Hughes, "On the mean accuracy of statistical pattern recognizers", *IEEE Trans. Inform. Theory*, vol. 14, pp. 55-63, 1968.

[3]   L. Dong, *Classification Method and Application of Hyperion Image Forest Vegetation*. Jilin University, China 2013.

[4]   S. Zhou, "Application of Ant Colony Algorithm in Dimensionality Reduction and Classification of Hyperspectral Images". *Harbin Institute of Technology*: China, 2010.

[5]   J. Wang, and G. Liu, "MaHaitao, Optimal band combination in macro-monitoring", *China Sci. Technol. Inform.,* vol. 16, pp. 40-41, 2011.

[6]   J. Li, Y. Xing, and C. Yang, "Optimal band selection based on the spectral characteristics of forest types-for example of HJ/lA hyperspectral images", *Forest Eng.*, vol. 29, pp. 4-5, 2013.

[7]   M.A. Na, H.Y. Fang, Z. da Feng, and W.X. Sheng, "Determination on the optimum band combination of hj-1a hyperspectral data in the case region of Dongguan based on optimum index factor and J-M distance," *Remote Sens. Technol. Appl.*, vol. 25, pp. 359-362, 2010.

[8]   Z. Wang, "Application of Hyperspectral Remotesensing in Identification of Forest Tree Species", *Zhejiang Forestry University*: China, 2011.

[9]   J. Li, "Identification of tree species based on environmental and disaster monitoring and forecasting satellites", *J. Northeast Forestry Univ.*, vol. 11, pp. 42-43, 2013.

[10]   R. Han, *Key technology of environmental satellite HSI image processing*. Central South University: Changsha, 2010.

[11]   S. Chu, "Computational intelligence based on the behavior of cats", *Int. J. Innovat. Comput. Inform. Control.*, vol. 3, pp. 163-173, 2007.

[12]   S. C. Chu, P. Tsai, and J. Pan, "Cat Swarm Optimization," In: *Proceedings of the 9th Pacific Rim International Conference on Artificial Intelligence,* China, pp. 854- 858, 2006.

[13]   G. Panda, P. M. Pradhan, and B. Majhi, "IIR System Identification Using Cat Swarm Optimization", *Expert Syst. Appl.*, vol. 10, pp. 12671-12683, 2011.

[14]   G. Wang, S. Yang, and F. Feng, "Image classification based on cat swarm algorithm study", *J. Tianjin Univ. Technol.*, vol. 27, pp. 35-39, 2011.

[15]   Q. Du, and Y. He, "Similarity-based unsupervised band selection for hyperspectral image analysis," *Geosci. Remote Sens. Lett.*, vol. 5, pp. 564-568, 2008.

[16]   Z.H. Wang, C.C. Chang, and M.C. Li, "Optimizing least significant bit substitution using cat swarm optimization strategy", *Inform. Sci.*, vol. 192, pp. 98-108, 2012.