International Journal of Signal Processing, Image Processing and Pattern Recognition
Vol. 8, No. 4 (2015), pp. 325-338
http://dx.doi.org/10.14257/ijsip.2015.8.4.28

A Band Selection Method for Hyperspectral Image Classification based on Improved Particle Swarm Optimization

Jie Shen[1,2], Chao Wang[1], Ruili Wang[3], Fengchen Huang[1], Chao Fan[1,4] and Lizhong Xu[1*]

[1]*College of Computer and Information Engineering, HoHai University, Jiangsu Nanjing 210098, China*
[2]*College of communication Engineering, PLA University of Science and Technology, Jiangsu Nanjing 210098, China*
[3]*School of Engineering and Advanced Technology, Massey University, Auckland, New Zealand*
[4]*7220 mailbox Beijing, China*
**Lizhxu@hhu.edu.cn; Shenjie_2003045@hhu.edu.cn*

Abstract

With the development of spectral imaging technology, it makes hyperspectral imagery widely used. According to the features of multiple bands and the strong mutual correlation among these bands, this paper presents a band selection method for hyperspectral imagery classification based on improved PSO (Particle Swarm Optimization). First of all, we use information divergence to describe the correlation of the bands, then build the information divergence matrix to make the classification of subspaces. Secondly, we construct the fitness function of the algorithm with the band information and categories of the Bhattacharyya distance (B distance) to improve the inertia weight updating method in PSO. Finally, based on the AVIRIS hyperspectral imagery and compared with existing method to conduct experiments, the average classification accuracy of the proposed method is 81.36%, which is distinctly improved 0.91% compared with the existed method. Meanwhile, the proposed method has a significantly faster convergence speed during the process of the band selection. Therefore, the experimental results verify the effectiveness of the proposed method in this paper.

Keywords: *Hyperspectral imagery, information divergence, PSO, band selection*

1. Introduction

Hyperspectral imagery technology as a kind of new earth observation technology has been widely applied in many fields [1-4]. At the same time, the large data volume and high redundancy take great difficulties for image processing. Reducing the dimension of the hyperspectral imagery is one of the effective ways to solve these problems.

The existing dimension reduction methods mainly include two categories, the feature selection method and the feature extraction method [5]. Feature selection method is a dimension reduction method based on non-transformation, which overcomes the drawback of transformation method to change the image characteristic that is helpful for maintaining the original image features. This paper mainly focuses on feature selection methods. At present, according to different applications, band selection method is mainly for four aspects in consideration. (1) Information theory angle, the information of the selected band should be as large as possible. (2) Statistically angle, the redundancy

* Corresponding Author

ISSN: 2005-4254 IJSIP

between the selected bands should be as little as possible. (3) Spectroscopy angle, the selected band should be with better spectral characteristics. (4) Classification angle, the separability of the selected band should be strong, it can get better classification accuracy [6-7]. With the less number of wavelengths, the information theory criterion can obtain better effect. If the correlations between adjacent bands are strong and the correlations between interval bands are weak, these selected bands are not typical from statistical point of view.

Divided subspace can effectively remove the inter band redundancy. It is beneficial for keeping global statistical properties and local statistical characteristics of image. At the same time, it can improve the efficiency of band selection, which is an important step of band selection [8-9]. Existing methods mainly depend on inter band correlation coefficient. Correlation coefficient is a linear relationship and when the two bands are the same or linear expression the correlation coefficient is the largest. In recent years, information divergence is used to represent the images correlation. Meer etc. compared the information divergence and correlation coefficient with the distinguish probability of spectrum and the ability probability of spectrum. The results indicated that information divergence is the best measure to describe the similarity of the images [10].

PSO [11] is an intelligent algorithm proposed by Kennedy and Eberhart which is inspired by birds foraging behavior. It is a parallel algorithm which can deal with a number of particles at the same time. It is easy to implement without complex iterative process of encoding, the final results have nothing to do with the selection of the initial state. Inertia weight decision the search ability of the algorithm. Shi and Eberhart believe that the bigger inertia weight can enhance the global search ability of the algorithm and the smaller inertia weight can enhance the local search ability of the algorithm [12]. Paper [13] had confirmed that when the inertia weight is greater than 1.2, the algorithm may fall into local optimum. Combined the above two ideas, in order to balance the global search and the local search ability by changing the inertia weight linear decreasing from the maximum 1.2 to minimum 0.1.

To solve the above problems and limitations, an improved PSO band selection method combining multiple criteria has been proposed in this paper. The method implements the improvements mainly based on the double rules including theory of information and separability between classes. On this basis, the method proposes an adaptability measurement which is composed of sum of entropies, variance of permutation entropy and Bhattacharyya distances between classes. The proposed adaptability measurement can enhance convergence speed and improve optimization quality effectively. Firstly, we described the correlation and divided the subspace with information divergence matrix. Secondly, the algorithm fitness function was build with the information of the band, the B distance of the two categories and the variance of the amount of information. The band combinations were selected by the improved PSO. Finally, experiments were conducted with the AVIRIS hyperspectral imagery and the existing methods are used for comparison. The selection bands were classification with minimum distance classification.

This paper consists four parts. The next chapter expounds the key technology involved and the specific implementation of the proposed method. The third chapter puts forward the method to make use of AVIRIS hyperspectral imagery for analysis and validation. The last chapter gives the conclusion.

2. Method

In this paper, band selection method for hyperspectral imagery classification based on improved PSO mainly concludes three stages. Firstly, it calculates the hyperspectral imagery spectral information divergence matrix. Secondly, the proposed subspace division method is used to delimit the molecular space. Finally, it selects the band

combinations with improved PSO. The implementation process of the proposed method is shown in Figure 1.

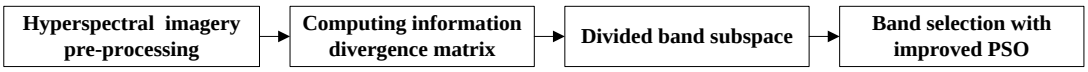

Figure 1. Data Processing Procedures

2.1. Information Divergence Matrix Calculation

The concept of information divergence comes from information theory [14]. It is firstly used to describe the difference between two probability distributions. It is used to measure the number of additional bit required to encode the sample P by the encoding of sample Q. In recent years, it has been used to indicate the correlation between the two images [15-17]. Each band of the hyperspectral imagery is interpreted as an image. Information divergence can be used to describe the correlation between two bands.

Information divergence is known as KL distance or relative entropy, which is the best measure of statistical independence. It can be used to describe the distance and difference between the two probability distributions. Assume that p and q are the two probability distributions both in the same alphabet Ω, information divergence can be used to characterize the different between them. Information divergence is defined as [18]:

$$D_{xy} = \sum_{\Omega} p(x) \log \frac{p(x)}{q(x)} + \sum_{\Omega} q(x) \log \frac{q(x)}{p(x)} \quad (1)$$

Where $p(x)$ represents the p probability of x , $q(x)$ represents the q probability of x . Set when $c > 0, c \log \frac{c}{0} = \infty$ .When $p(x) = q(x)$, the value is 0.

Although the information divergence is called distance, it is not a real sense of distance. In the traditional sense, the distance is required to meet: (1) Non-negative; (2) Symmetry; (3) The triangle inequality. Information divergence only meets the first. This paper used symmetric information divergence as a measure of similarity between two bands. The information divergence matrix of band is calculated according to equation (1).

2.2. Subspace Classification Method based on Information Divergence

In the process of hyperspectral imagery band selection, the basic principle of subspace division is to divide the strong correlation band into a set. By subspace division it selects one band from each subspace that can effectively remove the inter band redundancy. Information divergence as the measure of the band correlation that the bigger the value the weaker correlation between the bands. The principle of subspace division is to divide the number of continuous band in a group. The demarcation line is the local maximum points between the adjacent band information divergence curves.

2.3. Band Selection Method based on Improved PSO

PSO conducts the complex intelligent search by evaluation, comparison and simulate [19]. Each of individual in the algorithm can be seen as a particle without quality or volume. When flying, they adjust their speed and direction with their own or the group flying experience and the state. They gradually move closer to the optimal solution in the solution space. The merits of the fitness function are used to represent the particle. At the beginning of the algorithm, a number of particles are initialized randomly. Particles find the optimal solution by successive search. In the process of iterative the particles updating mainly depend on the extremum of the population and the existing optimal value. The velocity and position updating formula of particle is defined as [20]:

$$V_{id}^{k+1} = \omega V_{id}^{k} + c_1 r_1 (P_{id}^{k} - X_{id}^{k}) + c_2 r_2 (P_{gd}^{k} - X_{id}^{k}) \quad (2)$$

$$V_{id}^{k+1} = \begin{cases} V_{\max} & , \ V_{id}^{K+1} > V_{\max} \\ V_{\min} & , \ V_{id}^{K+1} < V_{\min} \end{cases} \quad (3)$$

$$X_{id}^{k+1} = X_{id}^{k} + V_{id}^{k+1} \quad (4)$$

Where V_{id}^{k+1} is the speed of the particles after iteration k + 1 times. ω is the inertia coefficient. c_1 and c_2 are the accelerating factors. r_1 and r_2 are the random numbers between 0 and 1. $V_{\max}$ and $V_{\min}$ are the maximum and minimum speed. X_{id}^{k+1} is the particle location after iteration k +1 times.

The speed of the particle after iteration k +1 times is influenced by its own history and the flying experience of groups. At the same time it is affected by the maximum flying speed and minimum flying speed.

In order to make the algorithm with better global search capability at the beginning of search and with better local search capability at the end of search, we made the variable inertia weight approach to search. Inertial weights with the method of linear transformation, the early stage of the iterations of inertia weight is bigger, which means the algorithm has better global search capability, while the late iteration inertia weight decreases, which means the algorithm has good local search ability to achieve a global search algorithm balance and local search capabilities. The transform formula of inertia weight ω is:

$$\omega = \omega_{\max} - \frac{(\omega_{\max} - \omega_{\min}) \times iter}{iter_{\max}} \quad (5)$$

Where ω is the inertia coefficient, $\omega_{\max}$ and $\omega_{\min}$ are the maximum and the minimum of inertia weight. $iter$ and $iter_{\max}$ are the current iteration number and the maximum number of the iterations. Suppose the maximum and minimum values of the inertia weight are 1.2 and 0.1 respectively. The maximum number of iterations is 500. Inertia weight value linearly decreases, and gradually decreases from the maximum value at the start of the search to a minimum

Each particle in the PSO represents a band combination. The band combination is determined by the fitness value of algorithm. In this paper, in order to overcome the one sidedness of the single criterion, respectively, we built the fitness function with the sum of band information entropy, band information entropy variance and B distance between two categories.

(1) The information entropy of the band i in the Hyperspectral imagery can be expressed as [21]:

$$H(x_i) = -\sum_{i=0}^{n} P_j \log_2 P_j \quad (6)$$

Where P_j is the spectral brightness value of the probability of j, n is the spectral brightness progression.

(2)The variance of the band entropy can be expressed as :

$$s^2 = \frac{1}{n} \sum_{i=1}^{n} (H(x_i) - \overline{H(x)})^2 \quad (7)$$

Where n is the number of band combinations, $H(x_i)$ is the entropy of the band, $\overline{H(x)}$ is the average information entropy of the band.

(3)At the processing of calculation, B distance has taken into account the time statistics and secondary statistics, which can be see the best distance measure of two categories in

high dimensional space. The greater the B distance between the two categories the better the separability. The B distance between category a and b can be expressed as [22-23]:

$$D_{ab} = \frac{1}{8}(\mu_a - \mu_b)^T \left(\frac{\sum_a + \sum_b}{2} \right)^{-1} (\mu_a - \mu_b) + \frac{1}{2} \ln \left[\frac{\left| \frac{\sum_a + \sum_b}{2} \right|}{\left(|\sum_a| |\sum_b| \right)^{\frac{1}{2}}} \right] \quad (8)$$

Where μ_a and μ_b are the mean spectral brightness, $\sum_a$ and $\sum_b$ are the covariance matrix corresponding region.

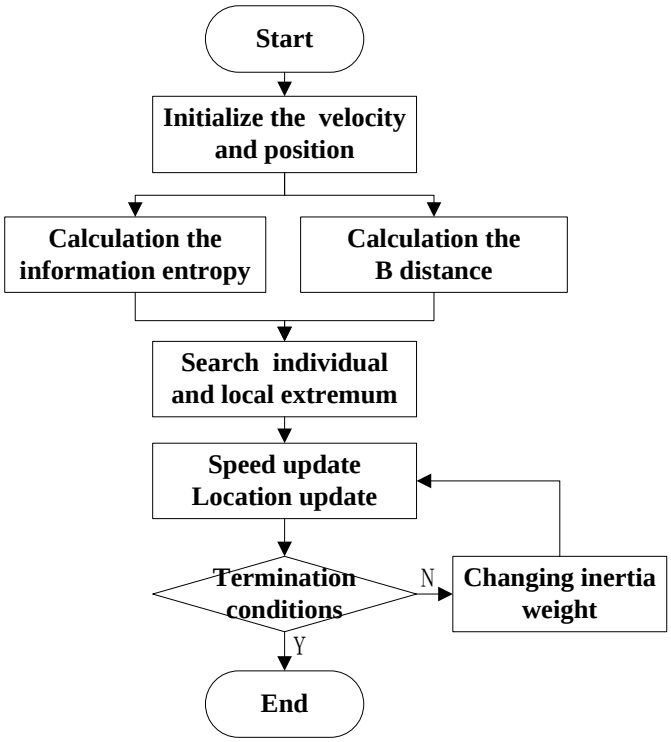

Figure 2. Band Selection Process

The band optimized selection process is shown in Figure 2 and the details of the selection step are shown as follows:

Step 1 To initialize the population particle of the algorithm and generate initial particles randomly.

Step 2 The particle fitness quality decides the selection probability of the particle. The fitness of particles is calculated according to the type (6-8).

Step 3 To compare the fitness value of particles in the population and to look for the qualified particle, then to conform to the conditions of particles in the solution set of inferior quality.

Step 4 To update the speed of particles according to the type (3) and to determine whether the particle velocity crosses the line according to the type (4).

Step 5 To update the particle's position according to type (5). In process of updating, whether it is more than the scope of the group is judged. If the new location is out of corresponding range it is operated with type (3) to avoid the new particle's position being out of band grouping.

Step 6 To update the best individual history location and group history position. Updating process can be divided into two steps. Firstly, the collection and new solutions in the old non inferior solution set is turned on. Secondly, in the collection of new solutions, mutual subjects are removed; new non inferior solution set is selected.

Step 7 To judge whether the iteration reaches the end condition, if not, go to step (4) to continue the loop iteration.

In the process of optimization, the extremum velocity and position of the particle is restricted to ensure the band combination coming from different subspace of band. The band combination selected by the method proposed in this paper contains more amount of information, lower correlation of bands, better separability between classes. The accuracy of classification can be improved effectively.

3. Experiment and Results

This paper used the Indiana agricultural regions in the United States acquired AVIRIS hyperspectral imagery (The original image 3 band synthetic false color images shown as Figure 3). This hyperspectral imagery has 220 bands, 145*145 pixels. Removed the band contaminated with moisture and imaging bands, the use of the remaining 179 bands to experiment. Experiments are on the computer which CPU is Intel (R) 2.10 GHz Core i3 and the memory is 2GB. The algorithm is implemented on Matlab R2009a.

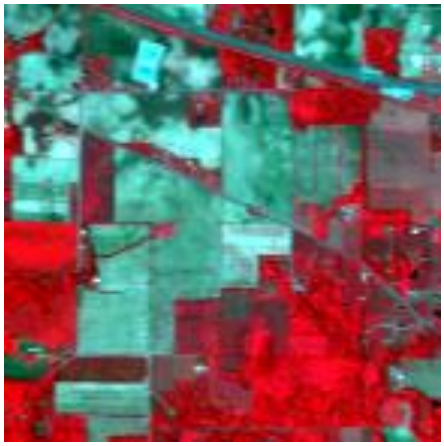

Figure 3. Synthesis of False Color Images

In order to illustrate the validity and feasibility of the band selection method in this paper, firstly, the subspace divide method proposed in this paper is compared with the method based on correlation coefficient. Secondly, the band selection method based on improved PSO is compared with that of PSO.

3.1. Results of Subspace Division

The image of information divergence grayscale and correlation coefficient matrix grayscale is shown in Figure 4 (a) (b). We can see by comparing the two Figures, the gray level in information divergence matrix is more abundant than that in the correlation coefficient grayscale. This explains the information divergence as the similarity measurement of band is better than correlation coefficient.

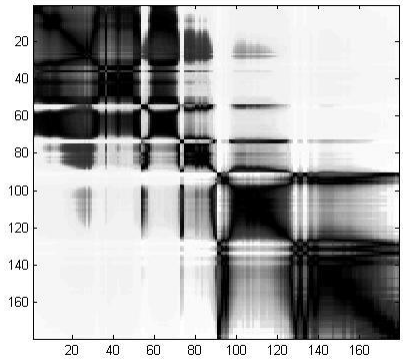

Figure 4(a). Information Divergence Matrix

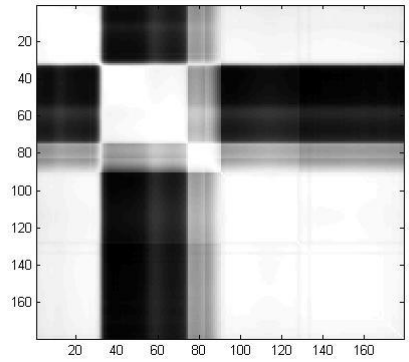

Figure 4(b). Correlation Coefficient Matrix

The curve of information divergence of the adjacent band is shown in Figure 5. Red tag is the local extremum points and divided subspace from local maximum value point. The subspace result is shown in Table 1.

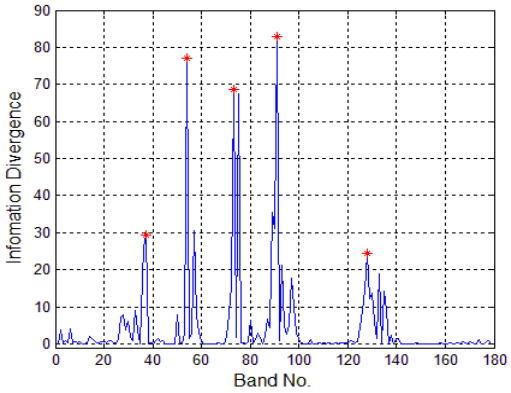

Figure 5. The Information Divergence between Adjacent Band Curve

Table 1. The Results of Subspace Classification

No.	1	2	3	4	5	6
Band	1-36	37-53	54-72	73-90	91-127	128-179
Dimension	36	17	19	18	37	52
ACC	0.8993	0.9944	0.9917	0.8938	0.9867	0.9682

In order to illustrate the correctness of demarcati subspace between the local maxima information divergence curve. The study focused on the analysis of before and after changes of local extreme point near the band and compared the band grayscale. The representative band 36, 37 are compared with the before and after of the band 127, 128. The 34-39 and 125-130 band grayscale are shown in Figure 6(a)(b). The band image quantization level is 256. The band grayscale outlines are roughly the same in Figure 6(a)(b). In image details grayscale similarity is higher among the 34, 35, 36 band in figure (a), as well the 37, 38, 39, band, meanwhile, there is an obvious difference between the grayscale of the 36, 37 band. The grayscale similarity is higher among the 125, 126, 127

band in Figure (b), as well the 128, 129, 130 band, meanwhile, there is an obvious difference between grayscale of the 128 and 127 band.

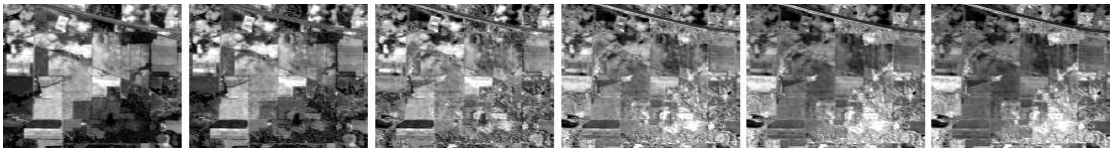

Figure 6(a). AVIRIS 34 to 39 Band Hyperspectral Remote Sensing Image Grayscale

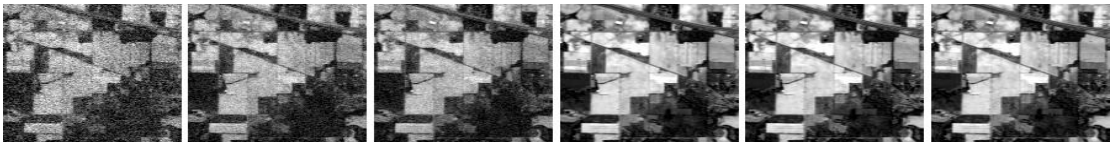

Figure 6(b). AVIRIS 125 to 130 Band Hyperspectral Remote Sensing Image Grayscale

In order to compare more intuitively, the ACC (average correlation coefficient) of subspace is used. The results of the subspace classification by the two methods are shown in Table 2. The dimension of the subspace divided by the method in this paper is less and the ACC of subspace is bigger. This shows that the inter band correlation in the subspace is strong.

3.2. Results of Band Selection and Classification

The method selects a typical band from each child subspace, so that each band combination contains 6 bands. The dimension of improved particle swarm algorithm population is set to 6. 50 particles are generated after each iteration. The max iteration time in the algorithm is 500. The inertia weight gradient linear decreases from 1.2 to 0.1 according the type (5). The initial particles of the algorithm are generated randomly, and each particle is a band combination.

Table 2. Subspace Classification Results

Method	No.	1	2	3	4	5	6
Method in ths paper	Band	1-36	37-53	54-72	73-90	91-127	128-179
	Dimension	36	17	19	18	37	52
	ACC	0.8993	0.9944	0.9917	0.8938	0.9867	0.9682
Comparative approach	Band	1-33	34-74	75-83	84-129	130-179	
	Dimension	33	41	9	46	50	
	ACC	0.9036	0.9647	0.9883	0.8988	0.9709	

B distance is the measure between two categories in the method which is proposed in this paper. B distance is used to characterize the two classes separability between A and B. The greater the B distance value is the better the separability between category A and B is. In the experiments, the principle of feature selection is according to the similar classes but not identical. The size of the samples selection in the image is 25 * 25 pixels (the total is 625 pixels).

Evaluating the pros and cons of a set of band combination is mainly depended on the amount of information, and the separability between classes. The information of the band combination is mainly considering how much information in each band of the combination. The more the information contained is the better the band combination is. The correlation between the band combination is mainly considering the correlation between the bands in the combination. The smaller the correlation between the bands is the better the band combination is. Separability is mainly reviewing the classification accuracy of band combination. The bigger the classification accuracy is the better the band combination is.

The band combination is selected from the improved PSO and standard PSO. The band combination selected from the two methods are in table 3. We respectively compare the band combination with the three aspects above.

(1) The average correlation coefficient

The smaller correlation is, the better it is. Each band combinations selected by the algorithm include six bands. An average correlation coefficient is introduced to compare the correlation in bands more directly. As shown in Figure 7, The figure in the abscissa is for the corresponding band combination, and the vertical axis is for the ACC corresponding band combinations. This paper presents a method to improve the use of selected band ACC coefficient which is smaller than that of standard algorithms ACC. It shows the less correlation and redundant inter.

(2) Separability between classes

B distance is to characterize the merits of separability between the two categories. This paper selects B distance as the parameter. The lager value of B distance means the better separability of combination bands. Figure 8 is the chart of a combination of B-band distance comparison. The graph shows that B distance of each band selected by the algorithm is better.

Table 3. The Select Band Combination

Method	No.	Band combination	Entropy	B distance
Method in ths paper	1	33 37 69 76 111 141	5.29	1.03
	2	33 37 69 77 113 141	5.22	1.03
	3	31 37 69 82 112 142	5.21	1.02
	4	14 36 65 77 112 142	5.24	0.92
	5	33 38 68 78 99 151	5.20	0.90
Comparative approach	1	35 38 64 77 102 143	5.25	1.02
	2	25 37 65 78 103 144	5.21	1.02
	3	24 39 66 76 101 146	5.25	0.99
	4	22 40 63 79 104 145	5.20	0.90
	5	21 44 62 75 100 142	5.18	0.90

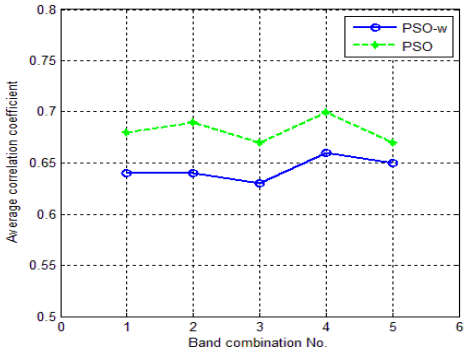

Figure 7. The Average Correlation Coefficient Chart between Band Combination

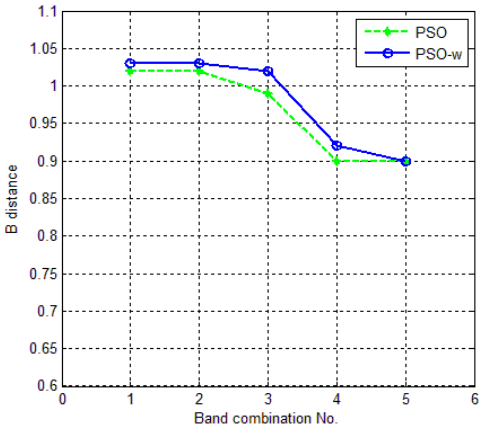

Figure 8. The B Distance Comparison Chart between Band Combination

Those are easy to distinguish between classes are able to improve the accuracy of classification.

(3) Category verification

Real experimental images feature a total of 16 different categories, nine of which has a small proportion of the entire image. Thus, only remaining seven images are taken into consideration. The number of feature classification category is set to 7 and the training samples, the classification of the test samples are selected in accordance with the ratio of 1:2. The number of test samples is showed in Table 5.

10 band combinations selected by two different groups are classified and compared in the standard of classification accuracy, 10 group classification accuracy for each band are shown in Table 6 and 7. Band selection algorithm using improved selection is compared with contrast band combination of high precision classification algorithm. Classification accuracy is in descending order and it reflects such an accuracy approximation. The overall classification accuracy of two algorithms are listed in Figure 9. The overall classification accuracy by using the improved algorithm to select band combination is higher than that of the standard algorithm. Original image feature comparison chart and band combinations classification results are shown in 9 (a) - (l) below.

Table 5. The Training Sample, Test Sample Size Classification Experiment

Category	Category 1	Category 2	Category 3	Category 4	Category 5	Category 6	Catagory7
Training sample	105	84	140	133	389	96	195
Test sample	210	168	280	266	778	192	390

Table 6. Improved PSO Selected Band Combination Classification Accuracy

No.	Category 1	Category 2	Category 3	Category 4	Category 5	Category 6	Catagory7	accuracy
1	68.09	83.33	92.14	78.20	89.72	71.87	84.10	83.76
2	66.19	86.90	86.78	81.20	84.70	80.73	79.74	81.83
3	58.10	85.11	78.57	82.33	94.47	62.50	76.15	81.26
4	88.10	64.89	78.21	70.31	82.00	76.56	88.72	80.17
5	85.24	58.92	85.71	72.18	80.20	67.71	92.05	79.82

Table 7. PSO Selected Band Combination Classification Accuracy

No.	Category 1	Category 2	Category 3	Category 4	Category 5	Category 6	Catagory7	accuracy
1	63.33	71.43	90.71	82.70	88.69	69.79	84.10	82.27
2	61.43	89.88	78.92	92.48	84.83	75.52	76.92	81.09
3	89.52	58.93	71.78	78.57	91.26	63.02	76.41	79.95
4	69.52	72.62	92.85	54.14	87.40	77.08	83.08	79.86
5	85.71	48.80	75.71	81.20	79.53	56.77	84.10	79.12

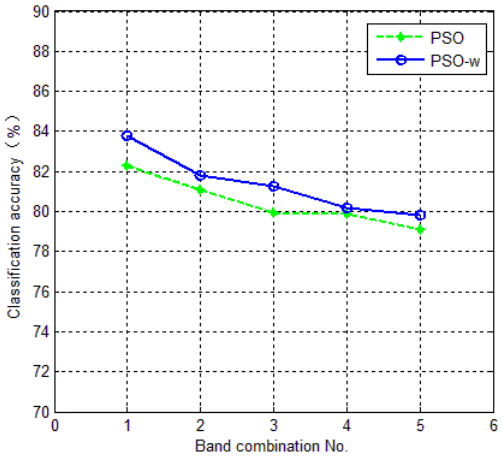

Figure 9. Band Combination Classification Accuracy Comparison Chart

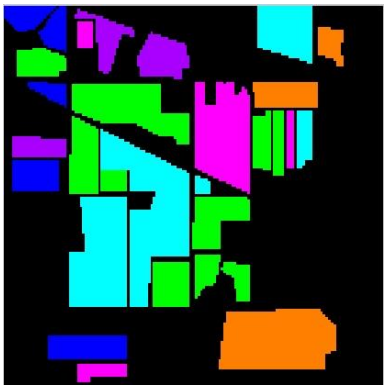

Figure 10(a). Original Feature Comparison Chart

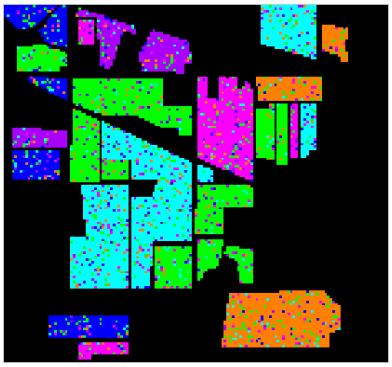

Figure 10(b). The Classification Result

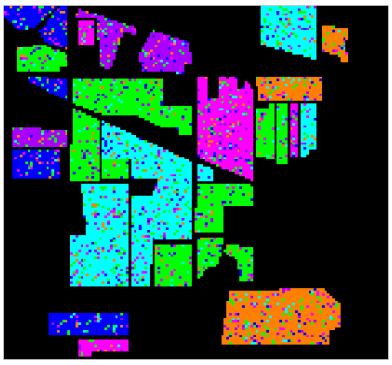

Figure (10). The Classification Result

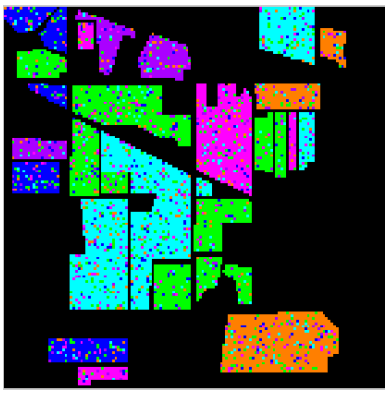

Figure 10(d). The Classification Result

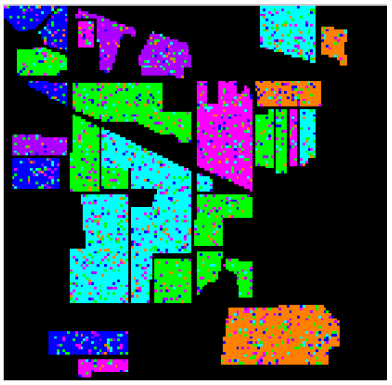

Figure (10). The Classification Result

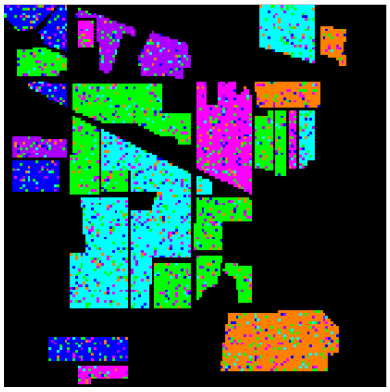

Figure 10(f). The Classification Result

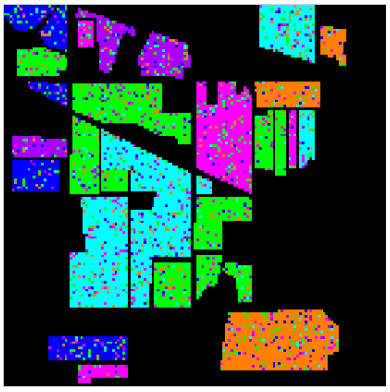

Figure 10(g). The Classification Result

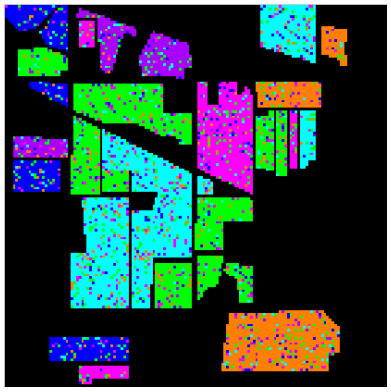

Figure 10(h). The Classification result

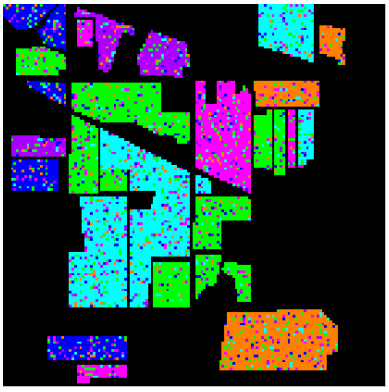

Figure 10(i). The Classification Result

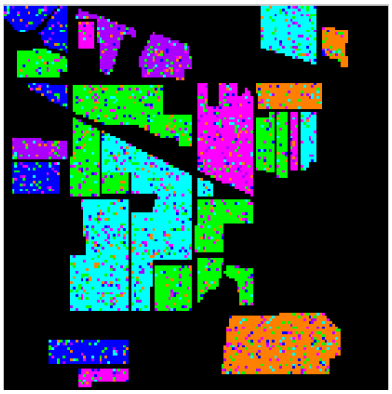

Figure 10(j). The Classification Result

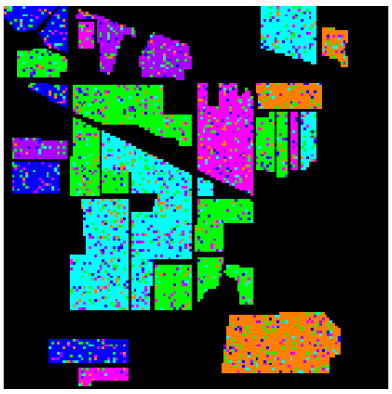

Figure 10(k). The Classification Result

For the improved algorithm with variable inertia weight approach, the global search ability and local search capabilities are both taken into account. Fitness value curve of the two algorithms are shown in Figure 11. It is obvious that improving the convergence rate significantly is faster than the standard algorithm and has a better search results.

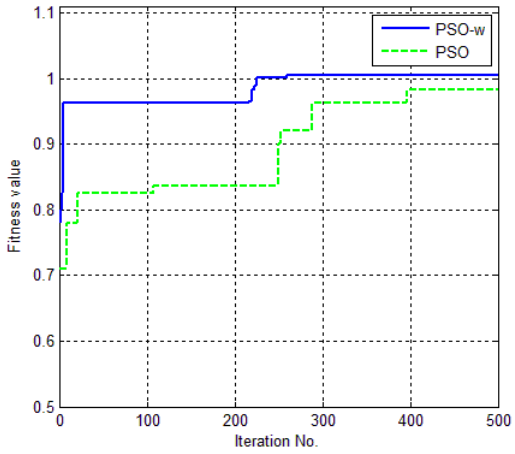

Figure 11.The Curves of Two Kinds of Algorithm Fitness

4. Conclusion

In this paper, in view of the one sidedness and limitation of existing dimension reduction method, information divergence characterization of inter band correlation is used to divide band subspace which overcomes the one sidedness of using partial correlation coefficient. In this paper a band selection method is proposed for hyperspectral imagery classification based on improved PSO (Particle Swarm Optimization). The method is considered from three aspects of information theory, statistics and image classification. Then the sum of the entropy of the band, the band combination entropy variance and the category B distance are chosen as band selection of fitness. It shows that the method of band selection algorithm significantly accelerates the convergence speed and has a better band combination selection. By experiments in this paper, the band combination of the average distance between correlation coefficient and category B selected by this method have higher classification accuracy than the standard particle swarm algorithm, and significantly accelerate the convergence speed of the algorithm.

References

[1] J. N. Price, I. Hiiesalu, P. Gerhold and M. Partel, "Small-scale grassland assembly patterns differ above and below the soil surface", Ecology, vol. 93, no. 6, **(2012)**.
[2] A. Y. Shi, G. B. Wu and D. B. Tan, "Multispectral remotely sensed image change detection based on EM algorithm and MDL rule", Journal of Hohai University (Natural Sciences), no. 2, **(2013)**, pp. 171-176.
[3] Q. X. Tong, B. Zhang and L. F. Zhen, Editor, "The interdisciplinary application of hyperspectral remote sensing", Publishing House of Electronics industry, Beijing, **(2006)**.
[4] L. P. Zhang and L. F. Zhang, Editor, "Hyperspectral Remote Sensing image", Publishing House of Wuhan University, Wuhan, **(2005)**.
[5] S. G. Bajwa, P. Bajcsy and P. Groves, "Hyperspectral image data mining for band selection in agricultural applications", Transactions-american society of agricultural engineers, vol. 47, no. 3, **(2004)**.
[6] J. Ling, L. Z. Xu and A. Y. Shi, "Remote sensing images fusion based on wavelet coefficients selection using choquet fuzzy integral", Journal of Remote Sensing, no. 2, **(2009)**, pp. 263-268.
[7] X. Wang, X. J. Yan, G. F. Lv and T. H. Fan, "Balloon-borne spectrum–polarization imaging for river surface velocimetry under extreme conditions", Infrared Physics & Technology, vol. 58, **(2013)**.
[8] Z. Shuang, "Research on the applicaition of ant colony algorithm in the dimentionality reduction and classification for hyperspectral image", Harbin Institute of Technology.
[9] X. J. Yan, Z. Zhang and Z. Chen, "FHT-CC-based adaptive motion vector estimation method for flow field image", Chinese Journal of Scientific Instrument, vol. 35, no. 1, **(2014)**.
[10] F. van der Meer, "The effectiveness of spectral similarity measures for the analysis of hyperspectral imagery", International journal of applied earth observation and geo information, vol. 8, no. 1, **(2006)**.
[11] J. Kennedy and R. C. Eberhart, Editors, "Particle Swarm Optimization", Proceedings of IEEE International Conference on Neural Networks, **(1995)** November 27-December 1, Perth, Western Australia.
[12] A. Ratnaweera, S. Halgamuge and H. C. Watson, "Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients", Evolutionary Computation, IEEE Transactions on, vol. 8, no. 3, **(2004)**, pp. 240-255.
[13] L. Palafox, N. Noman and H. Iba, "Reverse engineering of gene regulatory networks using dissipative particle swarm optimization", Evolutionary Computation, IEEE Transactions on, vol. 17, no. 4, **(2013)**, pp. 577-587.
[14] H. Jiao, Y. Zhong and L. Zhang, "Artificial DNA computing-based spectral encoding and matching algorithm for hyperspectral remote sensing data", Geoscience and Remote Sensing, IEEE Transactions on, vol. 50, no. 10, **(2012)**, pp. 4085-4104.
[15] M. Hasanlou and F. Samadzadegan, "Comparative study of intrinsic dimensionality estimation and dimension reduction techniques on hyperspectral images using K-NN classifier", Geoscience and Remote Sensing Letters, IEEE, vol. 9, no. 6, **(2012)**, pp. 1046-1050.
[16] M. X. Xu, F. Xu, C. R. Huang and M. Li, "Image restoration using majorization-minimizaiton algorithm based on generalized total variation", Journal of Image and Graphics, vol. 16, no. 7, **(2011)**, pp. 1317-1325.
[17] A. Y. Shi and M. Tang, "Super-resolution of remotely sensed images based on map and Discontinuity Adaptive markov random field", Intelligent Automation and Soft Computing, vol. 17, no. 7, **(2011)**, pp. 945-956.
[18] N. Paul and Y. Fan, "Angular discretization errors in transport theory", An information-based approach, Nuclear Science and Engineering, vol. 112, no. 3, **(1992)**, pp. 231-238.
[19] W. Zhang, "An image filtering method based on improved particle swarm optimization algorithm", International Review **on** Computers and Software, vol. 7, no. 3, **(2012)**, pp. 1405-1411.
[20] R. C. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory", Proceedings of the sixth international symposium on micro machine and human science, vol. 1, **(1995)**, pp. 39-43, Nagoya, Japan.
[21] T. M. Cover and J. A. Thomas, "Elements of information theory", John Wiley & Sons, **(2012)**.
[22] H. B. Wang, Z. Chen, X. Wang and Y. Ma, "Random finite sets based UPF-CPHD multi-object tracking", Journal of China Institute of Communications, vol. 33, no. 12, **(2012)**, pp. 147-153.
[23] C. H. You, K. A. Lee and H. Li, "An SVM kernel with GMM-supervector based on the Bhattacharyya distance for speaker recognition", Signal Processing Letters, IEEE, vol. 16, no. 1, **(2009)**, pp. 49-52.

