www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

A HYPERSPECTRAL BAND SELECTION BASED ON GAME THEORY AND DIFFERENTIAL EVOLUTION ALGORITHM

Aiye Shi [1,2], Hongmin Gao[1], Zhenyu He[1], Min Li[1] and Lizhong Xu[1]*

[1] College of Computer and Information Engineering, Hohai University, Nanjing 211100, China

[2] College of IOT Engineering, Hohai University, Changzhou 213022, China

Emails: ayshi@hhu.edu.cn, *Corresponding Author: cshhunj@126.com

Submitted: March 18, 2016 ***Accepted: Oct.19, 2016*** ***Published: Dec.1, 2016***

Abstract- This paper uses the combination of information and class separability as a new evaluation criterion for hyperspectral imagery. Moreover, the correlation between bands is used as a constraint condition. The differential evolution algorithm is adopted during the search of optimal band combination. In addition, the game theory is introduced into the band selection to coordinate the potential conflict of searching the optimal band combination using information and class separability these two evaluation criteria. The experimental results show that the proposed method is effective.

Index terms: Remote Sensing, Hyperspectral imagery, band selection, game theory, differential evolution algorithm.

I. INTRODUCTION

Hyperspectral image can record the power of electromagnetic wave from the reflection, absorption and radiation of some region [1]. It usually contains dozens of or even hundreds of spectral bands, and the whole image can be seen as an image cube. Although hyperspectral image can provide the rich spectral information for the application of remote sensing, it also brings the difficulty in the analysis and processing of data because of a large amount of redundant information. Therefore, it is necessary to reduce the dimensions of hyperspectral image.

Band selection is an important method of dimension reduction of hyperspectral image. At present, band selection methods mainly use information or class separability as an evaluation criterion. Liang Ge *et al* adopted mutual information as an evaluation criterion and then select the bands of center of classes as the optimal band combination [2]. Xuesong Liu *et al* proposed an unsupervised band selection method where the K-L divergence is applied to represent the information [3]. Padma S *et al* adopted Jeffries Matusita distance to select band combination [4]. Chenming Li *et al* searched the optimal band combination with particle swarm optimization based on Bhattacharyya distance. The above band selection methods all adopted the single criterion to evaluate good or bad of the chosen band combination. For example, when the information is used as an evaluation criterion, the selected band combination has a large amount of information. But the statistical distance between target classes may not be the largest, that is, the class separability may not be best. In like manner, when the class separability is used as an evaluation criterion, the information of the selected band combination may not be the optimal and therefore influence the accuracy of classification.

Gurram P and Kwon H proposed a band selection method based on coalition game theory [6]. The method firstly partitions the whole bands into several band subsets and uses the shapely value of coalition game theory to compute the contribution degrees for classification of each band subset. Then the method selects the band subset by setting the threshold of contribution degree and uses the selected band subset to implement the dimension reduction of hyperspectral imagery. Because the method traverses as far as possible all band subset, the computational cost is large.

Considered the one-sidedness for single evaluation criterion of band selection, some researchers proposed band selection method based on the comprehensive of multiple evaluation criteria.

Liguo Wang *et al* proposed the band selection method based on the weighted criteria of J-M distance and the optimal index factor [7]. Hongmin Gao *et al* adopted Choquet fuzzy integral to fuse the information, correlation between bands, and class separability for bands selection [8].

Multi-criteria band selection can be seen as the optimal problem of multi-target, multiple evaluation criteria correspond to multiple target of optimization. Band selection is the optimal problem of band combination for searching for the optimum of multi-objective. The above multi-criteria band selection methods adopt the basic strategy of transforming multi-objective into single objective and use search algorithm based on a single objective to search the optimal band combination.

However, the optimal problem of multi-objective differs from single-objective, it has a distinctive feature: one objective to be optimized possibly existing conflict with the other objective, that is, when one objective is optimized, the optimization of the other objective may be affected and can be degraded. Therefore, we can't adopt the strategy of optimization problem of a single objective for multi-objective. We can adopt some method to coordinate the conflict between multi-objectives and make them all obtain the optimal value [9]. The work in [10], [11], and [12] proposed the optimal problem of multi-objective based on game theory, where the idea of game theory is used to solve the optimization problem of multi-person and multi-object constraint and different objective functions, respectively. Inspired on the above work, this paper proposes the game theory to synchronize the information and class separability, where the correlation between bands is used to be constraint condition. In proposed method, the differential evolution algorithm is applied for searching the optimal band combination.

II. PROPOSED BAND SELECTION ALGORITHM

In this section, we briefly introduce the principle of game theory and differential evolution algorithm. Then, we give multi-criteria evolution function of band selection. Finally, the whole implementation steps of the proposed algorithm are given.

a. Briefly introduction to game theory

Game theory is the study of strategic decision making. It refers to one person choosing his/her strategy based on the strategy of other participant and finally winning or obtaining his/her maximization interest in a two or more persons equal game. The basic elements of game theory are composed of three basic components: players, set of strategies and utility function.

- Players: Players are the decision-makers of game theory, which chooses his /her action to be in order to maximize his/her interest. Set **P**= {1, 2, 3, … , *n*} represents the set of *n* decision-makers involved in game.
- Set of strategies: Set of strategies make the action scheme of one player replying to other player. Define s_p as a specific strategy of player *p*, $S_p=\{s_p\}$ represents possible all chosen strategies by player *p*. Strategies combination is defined as $\mathbf{S}=(s_1, s_2, s_3,\ldots, s_n)$.
- Utility function: Utility function is the pay-off after player gambling at specific strategies combination, which can provide more reasonable decision-making for players continuing to take part in game. Suppose that u_p represents the utility function of player *p* whose final pay-off under the influence of *n* players action is defined as: $\mathbf{u_p}=\mathbf{u_p}(s_1,s_2,s_3,\ldots,s_n)$。 The pay-off of *n* players under a game is defined as $\mathbf{U}=(\mathbf{u_1},\mathbf{u_2},\ldots\mathbf{u_n})$.

The final aim of game theory is to make the game of multi-person reach a balance of profit. This paper takes band selection method based on synchronization of multi-criteria as an optimization problem of multi-objective. Because there exists conflict during the optimization of multi-object, the game theory is introduced into coordinate profit between multi-object to obtain the balance and finally find the optimal bands combination by making both the information and class separability optimal.

b. Briefly introduction to differential evolution (DE) algorithm

DE is an evolutionary algorithm for solving the Chebyshev polynomial problem[13]. The implementation of DE algorithm includes the following four steps:

Step 1: Inialization operation

The operation randomly produces an initial population. Suppose $\mathbf{X_k(t)} = [x_{k1}(t),x_{k2}(t),\ldots,x_{kD}(t)]$ represent the individual *k* of population in the *t*-th iteration, and D represent the dimension of the problem of research. The value of $\mathbf{X_k(t)}$ can be represented as follows:

$$x_{kl}(t) = rand(L_l, U_l) \qquad (1)$$

where $x_{kl}(t)$ represents the value of *l*-th dimension of $\mathbf{X_k(t)}$, L_l and U_l represent the upper and lower bound of $x_{kl}(t)$, and rand(L_l,U_l) returns a random value between L_l and U_l, whose distribution is uniform distribution.

Step 2: Variation operation

The operation produces a new variation of $\mathbf{X_k(t)}$. Firstly, three different individuals of the current population, $\mathbf{X_{r1}(t)}$、$\mathbf{X_{r2}(t)}$、$\mathbf{X_{r3}(t)}$, are picked up, and then the variant individual $\mathbf{Y_k(t)}$ is produced according to the following equation:

$$\mathbf{Y}_k(\mathbf{t}) = \mathbf{X_{r1}}(\mathbf{t}) + F\left(\mathbf{X_{r2}}(\mathbf{t}) - \mathbf{X_{r3}}(\mathbf{t})\right) \tag{2}$$

where F represent the scale factor, whose value is in the range 0.1 to 1.

Step 3: Crossover operation

The aim of crossover operation is to make the current individual $\mathbf{X_k(t)}$ and the variant individual $\mathbf{Y_k(t)}$ crossover, and produce the diversity of population individual. We randomly produce an integer $r \in$ [1,D], and then each dimension of individual vector can be operated as follows:

$$\mathrm{h}_{kl}(t+1) = \begin{cases} y_{kl}(t), rand(0,1) \le CR \ or \ r = l \\ x_{kl}(t), \text{otherwise} \end{cases} \tag{3}$$

where $x_{kl}(t)$ represents the value of l-th dimension of $\mathbf{X_k(t)}$, $y_{kl}(t)$ represents the value of l-th dimension of $\mathbf{Y_k(t)}$, $h_{kl}(t+1)$ represents the value of l-th dimension of $\mathbf{H_k(t)}$ which represents intermediate individual after crossover operation, CR represent the rate of crossover, whose value is in the range of [0 1], and integer k can ensure the value of some dimension of intermediate individual different from current individual.

Step 4: Selection operation

The operation judge whether the result of the crossover operation can be selected to attend the next iterative evolution. It can judge whether intermediate individual superior to current individual based on the fitness function f and the selection operation can be represented as follows:

$$\mathbf{X}_k(\mathbf{t+1}) = \begin{cases} \mathbf{H}_k(\mathbf{t+1}), f(\mathbf{H}_k(\mathbf{t+1})) > f(\mathbf{X}_k(\mathbf{t})) \\ \mathbf{X}_k(\mathbf{t}), \text{otherwise} \end{cases} \tag{4}$$

In the DE algorithm, a variation pattern can generally be expressed as DE/x/y/z, where x indicates the selection method of the base vector in the mutation operation, y indicates the number of difference vectors, and z refers to the crossover operation pattern. The variation pattern introduced here can be expressed as DE/rand/1/bin, where bin refers to the binomial crossover model.

c. Evaluation function

In this study, band selection is carried on based on information and class separability, and information entropy and Bhattacharyya distance are chosen as evaluation criteria.

(1) Information entropy

According to Shannon theory [14], the entropy can represent information, which can reflect the degree of information. For hyperspectral image, suppose its data bit level is L bits, and then the information entropy of i-th band can be represented as follows:

$$H(i) = -\sum_{i=0}^{2^L-1} P_i(r)\log_2 P_i(r) \qquad (5)$$

where $H(i)$ represent the information entropy of i-th band, i=1,2...,D, r represents the pixel value, $P_i(r)$ represents probability for gray value r of i-th band. In general, the greater the information entropy of image is, the more information it will contain. By computing the information entropy of hyperspectral image, we can choose the band combination whose information is greatest.

(2) Bhattacharyya distance between classes

Bhattacharyya distance [15] (B distance) represent class separability, the larger B distance between two classes, the more separability between this two classes. B distance considers both one order statistical variant and two order statistical variant. B distance is the best measure between two classed in the high measure space. B distance can be represented as follows:

$$D_{ab} = \frac{1}{8}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \left(\frac{\sum_1 + \sum_2}{2}\right)^{-1} (\boldsymbol{\mu}_a - \boldsymbol{\mu}_b) + \frac{1}{2}\ln\left[\left(\left|\frac{\sum_1 + \sum_2}{2}\right|\right) \Big/ \left(\left|\sum_1\right|\left|\sum_2\right|\right)^{\frac{1}{2}}\right] \qquad (6)$$

where $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$ represent spectral mean of the indicated area of two classes, and $\sum_1, \sum_2$ represent spectral variance of the indicated area of two classes, respectively.

d. Steps of band selection

From the perspective of synchronization of multiple evaluation criteria, we think band selection as the optimization problem of multi-objective. Because among multi-objective exist conflicts during the optimization, that is, the multiple objective can not reach optimization simultaneously. In this study, the game theory is introduced into the optimization in order to coordinate multi-objective and make each objective obtain the optimization as far as possible. Intelligent search algorithms have been widely used in the band selection of hyperspectral image. And the DE algorithm is among those intelligent search algorithms, which has the following features: a simple structure, powerful global search ability and fast convergence. For the large number of bands of hyperspectral image, the study applies a binary DE search algorithm for band selection [16]. The concrete steps of the proposed method are as follows:

Step 1: Subspace division

The main features of hyperspectral image are the large number of bands, and high correlation and redundancy between neighbor bands. If band selection is faced with all bands, then some key local characteristic will lose, thus the chosen band combination may not helpful for classification. The basic thinking of resolving above problem is that all bands are divided into several subspaces, and where band selection is implemented. There are many common sub-space division methods. In this study, adaptive subspace decomposition (ASD) method based on correlation is adopted [17], which first compute correlation coefficient $R_{i,j}$ between two bands, where i and j represent i-th and j-th band of hyperspectral imagery. The value of $R_{i,j}$ is in the range -1 to 1. The larger the absolute of correlation coefficient between bands is, the strong correlation between bands is. The definition of $R_{i,j}$ is as follows：

$$R_{i,j} = \frac{\sum_{k=1}^{n}(x_{ik}-\overline{x_i})(x_{jk}-\overline{x_j})}{\sqrt{\sum_{k=1}^{n}(x_{ik}-\overline{x_i})^2\sum_{k=1}^{n}(x_{jk}-\overline{x_j})^2}} \tag{7}$$

where x_{ik} and x_{jk} represent the k-th pixel of i-th and j-th bands of hyperspectral image respectively, n represents the total number of pixels in a band, and $\overline{x_i}$ and $\overline{x_j}$ represent the mean of i-th and j-th band image.

Because the correlation between bands of hyperspectral image has the feature of block-division, thus the continuous bands with high correlation can be grouped into a subspace. Then, the band selection can be implemented in each subspace and the chosen bands can be combined to reduce the correlation between bands.

Step 2: Initialization of population

Before searching the optimal band combination with DE algorithm, the initialization for population is needed. Suppose the number of individual is m, there are N subspaces based on Step1. Then, M bands are chosen in each subspace, and $N\times M$ bands construct the band combination and seen as individual of population. Under the constraint of subspace division, randomly initialize m individuals.

Step 3: Construction of external set

DE algorithm will find some elite solution of band combination during iterated searching. The elite solution is that corresponding information entropy and B distance of these band combinations are larger than other band combination. In this step, the external set is built to store these elite solutions which rejoin of the population to take part in an iterated search. The benefit of this operation is that band combination of elite solution and population of algorithm at current compete together to produce the next generation population and is helpful to keep good quality of population.

Step 4: Decision-making

Inspiration by the algorithm of [18] which adopted clone selection algorithm to resolve the optimization of preference multi-objective, this paper constructs a limited repeated game model based on target preference. In this study, we take information entropy and B distance as two players of the game. The two players carry out limited repeated game during the iteration of algorithm. Because game repeats many times, rational player must consider balance of present interest and longtime while selecting strategic action at current stage of game. Player may choose strategy of not maximizing his/her interest at current game and abandon the present interest in order to keep long time interest. Therefore, there exists cooperation and conflict among players during the game. Player makes a decision whether to choose cooperation or to choose confrontation based on the interest of the previous stage. The concrete model is as follows:

Suppose information entropy and B distance of the whole population represent as $(E_1(t),E_2(t),...,E_m(t))$ and $(B_1(t),B_2(t),\ldots,B_m(t))$, respectively, then the fitness matrix **FITS(t)** of the population at t-th iteration can be represented as:

$$\mathbf{FITS(t)} = \begin{bmatrix} E_1(t), E_2(t),..., E_m(t) \\ B_1(t), B_2(t),..., B_m(t) \end{bmatrix} \tag{8}$$

Then, a utility matrix is built, which represent interest of other players from a strategy one player adopts. Utility matrix can be represented as

$$\mathbf{U(t)} = \begin{bmatrix} u_{11}(t), u_{12}(t) \\ u_{21}(t), u_{22}(t) \end{bmatrix} \tag{9}$$

where $u_{pq}(t)$ represent benefit of the p-th player from the q-th player. The final object of each player is to search the maximization of interest. In the band selection, the maximum profit is to search the band combination with maximum information entropy or B distance.

At the beginning of game, player will probe to adopt some strategy, expect himself/herself and his/her opponent to possibly increase interest based on final strategic action. If other players' strategy action of previous stage makes his/her interest increase, then reward strategy is possibly chosen during the choice of strategy action of later stage. Otherwise, punishment strategy is chosen. Define the set of strategy $S=\{s_1,s_2\}$, and two players share one set of strategy, and s_1 represent reward, and s_2 represent punishment. We build a reward choice probability matrix $\mathbf{M_{PR}}$ **(t)**, which can be represented as:

$$M_{\mathbf{PR}}\mathbf{(t)} = \begin{bmatrix} P_{11}(t), P_{12}(t) \\ P_{21}(t), P_{22}(t) \end{bmatrix} \qquad (10)$$

where $P_{pq}(t)$ represent probability that the p-th player choosing the q-th player choosing reward strategy, correspondly, the probability of choosing a punishment strategy is $1\text{-}P_{pq}(t)$. Player q can adjust reward choice matrix for p-th player based on his/her interest. If $u_{pq}(t) - \frac{1}{2}\sum_{p=1}^{2} u_{pq}(t) > 0$, that is the final interest of player q from the strategy of player p is larger than that of strategy of player q, then player q will adjust $\mathbf{M_{PR}}$ **(t)** matrix, furthermore, he/she will increase reward choice matrix $P_{pq}(t)$ for player p; otherwise, player q decrease $P_{pq}(t)$.

Based on the above definition, a preference weight matrix **W(t)** is built to embody each player’s strategy action, which is represented as follows:

$$\mathbf{W(t)} = \begin{bmatrix} w_{11}(t), w_{12}(t) \\ w_{21}(t), w_{22}(t) \end{bmatrix} \qquad (11)$$

where $w_{pq}(t)$ represents the preference degree of player p for player q. If player p choose reward strategy for player q, then the value of $w_{pq}(t)$ will increase; otherwise, it will decrease. Considering information entropy and B distance possibly differ great, we adopt normalization in order to prevent the information of little value flood by information of large value. We normalize the fitness value of each objective between 0 and 1. Therefore, the value of multi-objective fitness function for individual k of population on player p mapping fitness function $F^i_k(t)$ can be represented as follows:

$$F_k^p(t) = w_{p1}E_k^{'}(t) + w_{p2}B_k^{'}(t) \qquad (12)$$

where $E'_k(t)$ is normalization information entropy, and $B'_k(t)$ is normalization B distance. After two players adopt strategy to finish a game, let mapping fitness function of two players do the selection operation of DE algorithm to produce subpopulation, and merge two subpopulations into the new population of next generation.

Step 5: Iteration of DE algorithm

DE algorithm carries out variant, crossover and selection operations for subpopulation from the decision-making of the game to produce a new subpopulation. After two subpopulations merged, the non-dominated sorting and local search algorithm [19] is adopted to compute non-inferior level of the individual new population, the band combination whose non-inferior level is higher is elite band combination and join into the external set.

Step 6: Inspection of end condition

Inspect whether the terminal condition is satisfied (for example, reaching the maximum iteration number), if terminal condition is satisfied the optimization stops and nondominant band combination of external set is outputted.

e. Implementation of the proposed method

The flowchart of proposed method is as follows:

(1) Preprocessing of original hyperspectral image, remove the disturbing band and preselect object type.

(2) Division of subspace and initialization of the population under the constraint of subspace.

(3) Build external set and store elite band combination.

(4) Compute non-inferior level of individual of population, and put the good individual whose non-inferior level is high into external set.

(5) Produce mapping fitness value of individual of population by decision of game.

(6) Search subpopulation from decision-making of game by iteration, and put subpopulation into new generation population.

(7) Inspect terminal condition is satisfied, if not, return (4).

(8) Output elite band combination of external set, i.e., best combination.

III. TRADITIONAL WAYS OF FIXING FAULT CURRENT

a. Description of experimental data

In order to validate effectiveness of the proposed method, this paper carried out experiments on a standard hyperspectral image. The remote sensing image is a subsection of northwestern agriculture and forestry mixed test area of Indiana State of America which was acquired by in June on AVIRIS sensor. The wavelength ranges 0.4 to 2.5μm, and the size of image is 145×145 pixel, and the spatial resolution is 25m. 179 bands are kept by removing bands which is severely polluted by water vapor and noise (bands 1-4, 78, 80-86, 103-110, 149-165, and 217-224). Figure 1 is false color image composed by bands 50, 27 and 17. Figure 2 is the ground truth of object.

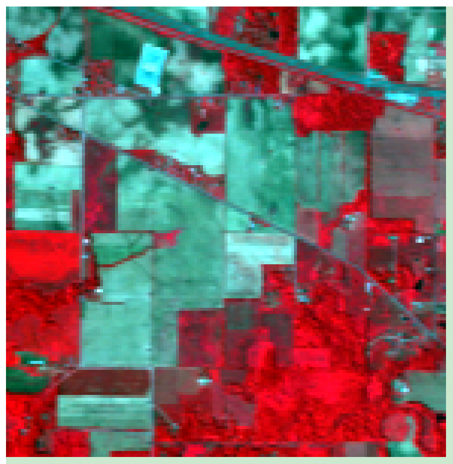


Figure 1 Color image composed by bands 50, 27, and 17

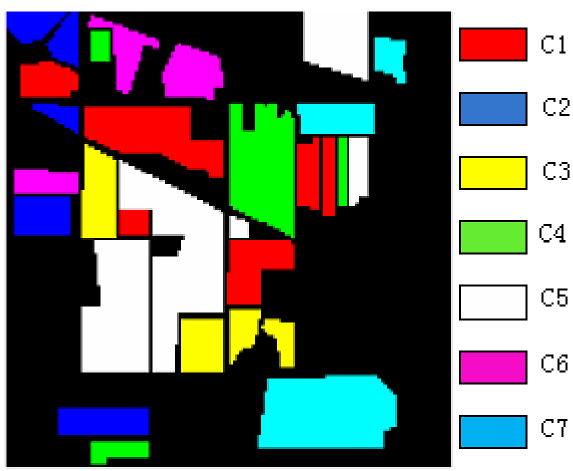


Figure 2 Ground truth

The research area includes 16 types of object and the samples of seven classes are chosen for experiments. Training samples and test samples are chosen uniformly according to the ratio 1:1.

Table 1 shows the class number, name, and the number of training and test samples. The program adopts software Matlab (R2009b) to implement. SVM classifier adopts LIBSVM toolbox (http://www.csie.ntu.edu.tw/~cjlin/libsvm/).

Table 1: Training samples and test samples

Number of class	Name of Classification	Training samples	Test samples
C1	Corn-notill	717	717
C2	Corn-mil	417	417
C3	Grass/Trees	374	373
C4	Soybeans-notil	484	484
C5	Soybeans-min	1234	1234
C6	Soybean-clean	307	307
C7	Woods	647	647
——	Total	4180	4179

b. Description of experimental data

We adopt adaptive subspace decomposition method based on correlation filtering for division of hyperspectral image. Matrix gray image of correlation coefficient (CC) is shown in Figure 3.

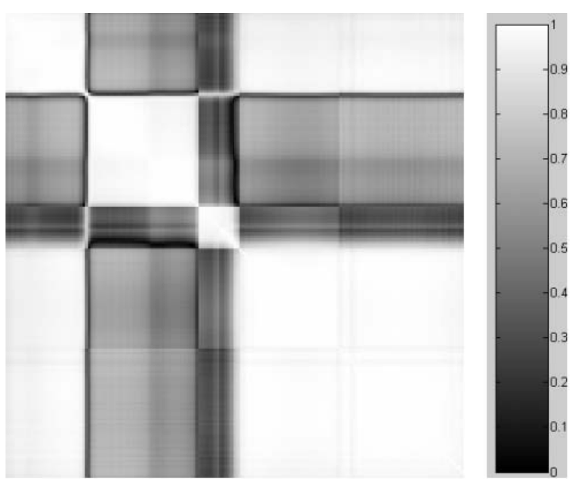


Figure 3 Matrix gray image of correlation coefficient

Lighter points in figure 3 are those of higher correlation coefficient. The value of CC equals 1 representing the lightest point and the value of CC of diagonal line of matrix. From figure 3, we can see that hyperspectral image has distinctive block-division features, and therefore, the subspace can reasonably be divided according to the CC between bands and the hyperspectral bands further can be in groups. Because CC matrix is large, Table 2 only gives CC data for partly bands (N represents the number of band).

Table 2: Matrix of correlation coefficient between partly bands

CC N \ N	20	21	22	23	24	25	26	27	28	29
20	1									
21	0.999	1								
22	0.999	0.999	1							
23	0.999	0.999	0.999	1						
24	0.997	0.996	0.996	0.997	1					
25	0.995	0.996	0.995	0.995	0.992	1				
26	0.994	0.993	0.992	0.992	0.995	0.993	1			
27	0.975	0.973	0.971	0.973	0.982	0.974	0.986	1		
28	0.890	0.886	0.881	0.886	0.905	0.888	0.917	0.957	1	
29	0.425	0.417	0.409	0.417	0.453	0.420	0.473	0.576	0.769	1

Suppose the number of subspace after decomposition is 5, and the dimension and the bands contained of each subspace are shown in Table 3.

Table 3: Dimension of subspace and bands contained

Number of subspace	1	2	3	4	5
Bands contained	5~35	36-76	77,79,87~97	98~102	111~148, 166-216
Number of bands	31	41	13	5	89

c. Experimental results and analysis

In this study, proposed method comprehensively considers three basic rules of band selection: information, class separability, and correlation between bands. In order to validate the proposed method effective, we compare the following band selection algorithms: information entropy based algorithm (IE), B distance based algorithm (BD) [5], synchronization of information entropy and class separability algorithm (SIEBD). SVM is chosen classifier to test the classification of band reduction.

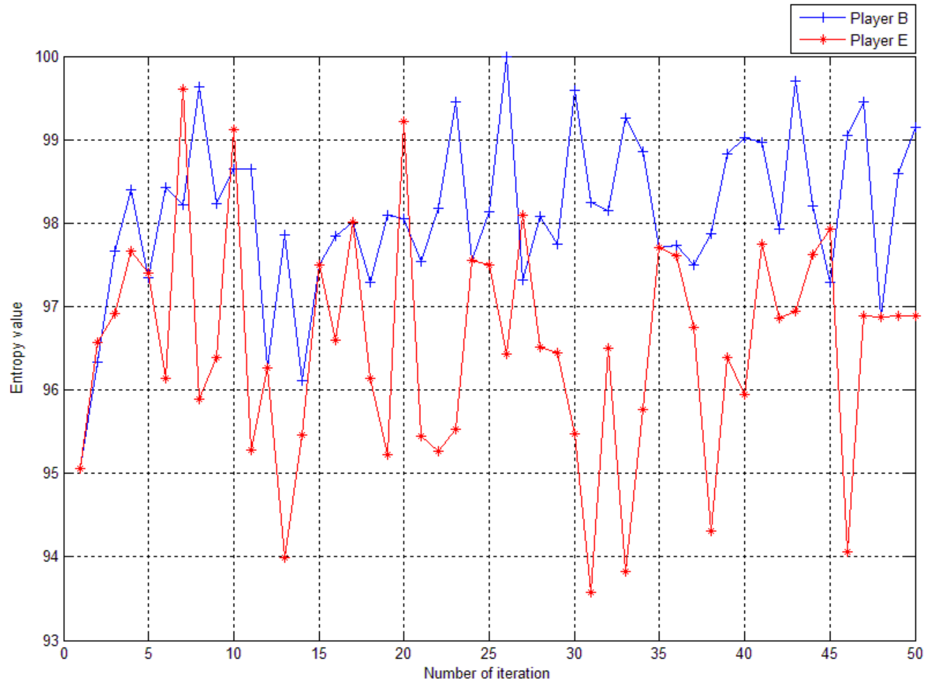
The relevant parameters of proposed method are set as follows: scale of population is 50; the dimension of individual of population is set to be D =10, which represents that randomly select 2 bands from 5 subspace (See Section 3.2) to construct 10 bands of band combination; the maximum iteration number is 50; the scale factor of DE algorithm is set to 0.6; the hybridization parameter is set to be 0.9; the variant model is set to be DE/rand/bin; RBF kernel is chosen in SVM classifier, and punishment parameter c and kernel parameter γ are chosen based on 5-fold cross validation (c = 16， γ= 2.2974).

Because more optimal solutions are obtained for multi-objective optimization problem, for each algorithm we extract 3 groups optimal combination solution for comparison. Table 4 shows the optimal band combination and the corresponding information entropy (B distance) and classification accuracy for each group of the compared algorithms.

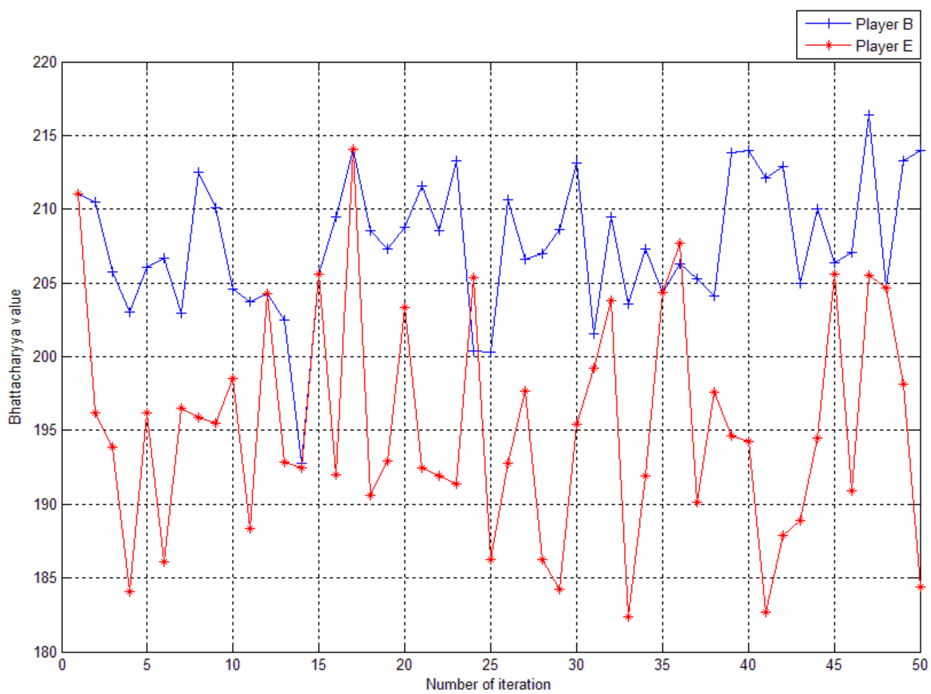
From experimental results of IE and BD algorithms, it can be seen that much information entropy or B distance May not helpful for classification accuracy. This demonstrates that a single evaluation criterion has one-sidedness. The results of SIEBD differ little from BD algorithm, and this shows that strategy with transferring multi-objective into single-objective may not be good to resolve optimization problem of multi-objective. Experiments of proposed algorithm show that the information entropy and B distance of band combination are not both maximal value but the results after the game with a compromise. Classification accuracy of three groups of proposed are high than other three groups of IE, BD, and SIEBD algorithms according average accuracy and maximal accuracy, which shows that the proposed method can search an effective band combination for classification. Figure 4 shows the processing of game, where player B adopts B distance evaluation criterion and player E adopts information entropy, respectively.

Table 4: Experimental results for 4 groups

Algorithm	**Band combination**	**Information entropy**	**B distance**	**Classification Accuracy**
IE	(21,31,40,52,87, 89,102,103,126,134)	95.2211		85.3571%
	(18,30,52,65,89, 92,102,111,119,122)	95.5217		83.8976%
	(20,24,49,52,89, 90,102,111,119,123)	96.5435		82.2586%
BD	(14,24,37,74,77, 91,102,103,133,180)		212.8078	88.8862%
	(20,28,44,70,92, 97,102,103,136,177)		216.0261	87.6540%
	(25,29,36,75,89, 92,103,111,136,181)		217.8357	87.2832%
SIEBD	(23,30,38,72,88, 89,102,111,136,184)	94.3385	219.1826	88.3718%
	(29,34,39,71,88, 90,102,103,134,198)	93.4926	214.7543	88.1326%
	(23,29,47,73,79, 88,102,103,136,181)	92.4287	219.2527	86.6850%
Proposed	(9,30,36,38,79, 89,98,99,137,201)	95.6491	192.7168	90.0825%
	(11,28,36,41,87, 91,101,102,131,181)	95.9316	211.4944	89.7356%
	(12,25,39,56,77, 90,98,100,143,186)	95.9765	179.5337	88.7427%



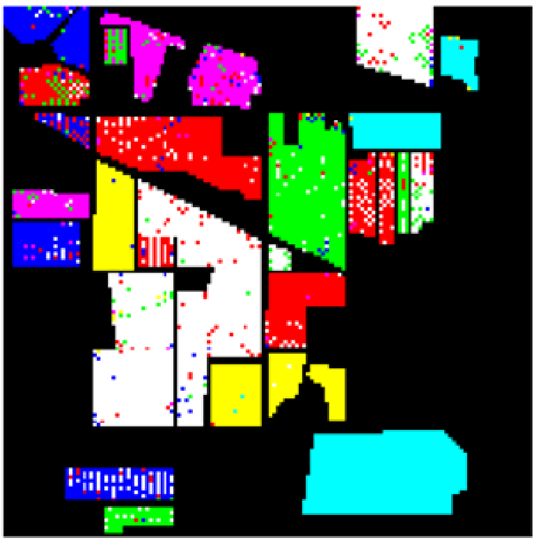
(a) Game processing of information entropy



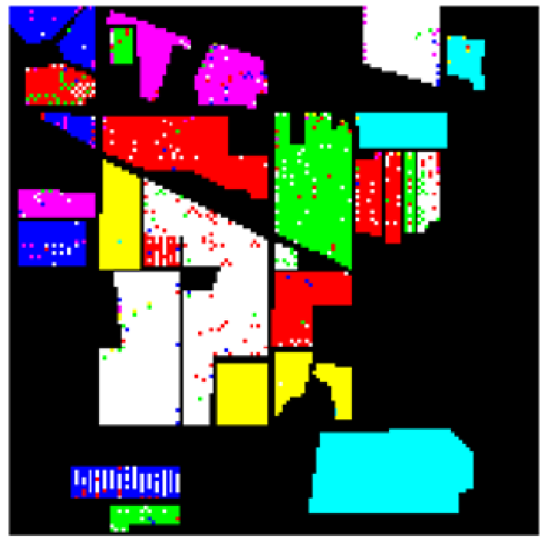
(b) Game processing of B distance

Figure 4 Game processing for information and entropy.

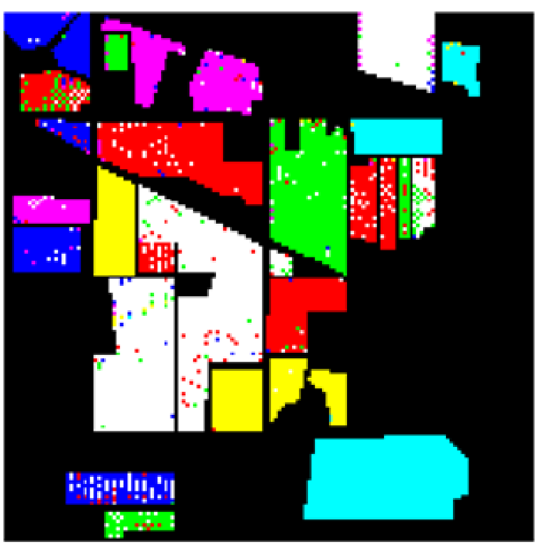
Figure 4 (a) and (b) shows that the curve of optimal individual based on the strategy under information entropy (player E) and B distance (player B) at each game processing, respectively. Figure 4 (a) shows the change of information entropy of optimal individual. Figure 4 (b) shows the change of B distance of optimal individual. Take Figure 5(a) as an example, at early stage of game, player E (information entropy) finds that the information entropy of optimal individual which is found by player B (B distance) is often slightly larger than that of his/her, this show that



(a) Result of IE method



(b) Result of BD method



(c) Result of SIEBD method

(d) Result of proposed method

Figure 5 Results of 4 group experiments

player B can help player E to find the individual with higher information entropy. Therefore, player E increases the preference of player B, i.e., chooses reward strategy and expects player B can search for him/her more optimal individual. In addition, player B per se searches the individual with larger B distance. In this case, each player thinks his own strategy is optimal, and therefore, the balance is obtained. For IE, BD, SIEBD and proposed algorithm, the classification results corresponding to the band combination with high class accuracy are shown in Figure 5.

IV. CONCLUSIONS

This paper proposes a new evaluation criterion for band selection for hyperspectral imagery. That is, the combination of information and class separability is used to be as a new evaluation criterion, at the same time, the correlation between bands is used as a constraint condition. Band selection of hyperspectral image is seen as multi-objective optimization problem. Differential evolution algorithm is used to search the band combination which can make multi-objective optimization. In addition, the game theory is introduced into the band selection to coordinate the potential conflict of multi-objective. The experimental results show that the proposed method can search better band combination than the methods based on the information, the methods based on class separability and the methods based on weighted information and class separability. The further optimization of the game theory in the band selection can be in the following two aspects: (1) probability of choosing reward strategy; (2) the weight of preference.

ACKNOWLEDGEMENT

This paper was supported by the National Natural Science Foundation of China (No. 61271386 and No. 41301448).

REFERENCES

[1] G.Shaw and D.Manolakis, “Signal processing for hyperspectral image exploitation”, IEEE Signal Processing,Magazine, vol. 19, no. 1, 2002, pp. 12-16.

[2] L.Ge, B.Wang and L.M. Zhang, "Band selection based on band clustering for hyperspectral imagery", Journal of Computer-Aided Design & Computer Graphics, vol. 24, no.11, 2012, pp. 1447-1454.

[3] X.S.Liu, L.Ge, B.Wang and L.M.Zhang, "An unsupervised band selection algorithm for hyperspectral imagery based on maximal information", Journal of Infrared and Millimeter Waves, vol. 31, no. 2, 2012, pp. 166-176.

[4] S.Padma and S.Sanjeevi, "Jeffries Matusita based Mixed-measure for Improved Spectral Matching in Hyperspectral Image Analysis", International Journal of Applied Earth Observation and Geoinformation, vol. 32, 2014, pp. 138-151.

[5] C.M.Li, Y.Wang, H.M.Gao and L.L.Zhang, "Band Selection for Hyperspectral Image Classification based on Improved Particle Swarm Optimization Algorithm", Advanced Materials Research, vol. 889-890, 2014, pp. 1073-1077.

[6] P.Gurram and H.Kwon, "Coalition Game Theory based Feature Subset Selection for Hyperspectral Image Classification", IEEE International Geoscience and Remote Sensing Symposium, Canada, 3446-3449, 2014.

[7] L.G.Wang and F.J.Wei, "Artificial physics optimization algorithm combined band selection for hyperspectral imagery", Journal of Harbin Institute of Technology, vol. 45, no. 9, 2013, pp. 100-106.

[8] H.M.Gao, L.Z. Xu, C.M. Li, A.Y. Shi, F.C. Huang and Z.L.Ma, "A New Feature Selection Method for Hyperspectral Image Classification based on Simulated Annealing Genetic Algorithm and Choquet Fuzzy Integral", Mathematical Problems in Engineering, 2013.

[9] Y.M. Meng, W.X. Li, Q.W. Chen, X. Yu, K.Y. Zheng and G.C. Lu "An Improved Multi-objective Evolutionary Optimization Algorithm for Sugar Cane Crystallization", International Journal on Smart Sensing and Intelligent Systems, vol. 9, No.2, 2016, pp.953-978.

[10] C.S. Lee, "Multi-objective Game-theory Models for Conflict Analysis in Reservoir Watershed Management", Chemosphere, Vol.87, no.6, 2012, pp. 608-613.

[11] P.Gurram and H.Kwon, "Coalition game theory based feature subset selection for hyperspectral image classification", IEEE International Geoscience and Remote Sensing Symposium, 3446-3449, 2014.

[12] M.Zamarripa, A.Aguirre, C.Mendez and A.Espuna, "Integration of Mathematical Programming and Game Theory for Supply Chain Planning Optimization in Multi-objective

Competitive Scenarios", 22nd European Symposium on Computer Aided Process Engineering, England, vol. 30, 2012, pp. 402-406.

[13] R.Storn and K.Price, "Differential evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces", Journal of Global Optimization, vol. 11, no. 4, 1997, pp. 341-359.

[14] K.M.Yang, S.W.Liu, L.W.Wang, J.Yang, Y.Y.Sun and D.D. He, "An Algorithm of Spectral Minimum Shannon Entropy on Extracting Endmember of Hyperspectral Image", Spectroscopy and Spectral Analysis, vol. 34, no. 8, 2014, pp. 2229-2233.

[15] T.Castaings, B.Waske, J.A.Benediktsson and J.Chanussot, "On the Influence of Feature Reduction for the Classification of Hyperspectral Images based on the Extended Morphological Profile", International Journal of Remote Sensing, vol. 31, no. 22, 2010, pp. 5921-5939.

[16] Y.C.Huo, X.Z.Wang and Y.Z.Kou, "A binary differential evolution algorithm with hybrid encoding", Journal of Computer Research and Development, vol. 44, no. 9, 2007,pp. 1476-1484.

[17] J.P.Zhang, Y.Zhang, B.Zou and T.X.Zhou, "Fusion classification of Hyperspectral Image based on Adaptive Subspace Decomposition", IEEE International Conference on Image Processing, Canada, vol. 3, 2000, pp. 472-475.

[18] D.D.Yang, L.C.Jiao, M.G.Gong and H.Yu, "Clone selection algorithm to solve preference multi-objective optimization", Journal of Software, vol. 21, no. 1, 2010, pp. 14-33.

[19] B.L.Chen, W.H.Zeng, Y.B.Lin and D.F.Zhang, "A New Local Search-Based Multiobjective Optimization Algorithm", IEEE Transactions on Evolutionary Computation, vol. 19, no. 1, 2015, pp. 50-73.