A band selection method of hyperspectral remote sensing based on particle frog leaping algorithm*

MU Lin-lin (穆琳琳)**, ZHANG Chao-zhu** (张朝柱)**, CHI Peng-fei** (池鹏飞)**, and LIU Lian** (刘潋)**

College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

(Received 27 February 2018; Revised 15 April 2018)

Dimensionality reduction is becoming an important problem in hyperspectral image classification. Band selection as an effective dimensionality reduction method has attracted more research interests. In this paper, a band selection method for hyperspectral remote sensing images based on subspace partition and particle frog leaping optimization algorithm is proposed. Three new evolution strategies are designed to form a probabilistic network extension structure to avoid local convergence. At the same time, the information entropy of the selected band subset is used as the weight of inter-class separability, and a new band selection criterion function is constructed. The simulation results show that the proposed algorithm has certain advantages over the existing similar algorithms in terms of classification accuracy and running time.

Document code: A **Article ID:** 1673-1905(2018)04-0316-4

DOI https://doi.org/10.1007/s11801-018-8028-7

The hyperspectral remote sensing data can be used to distinguish objects more accurately, but it makes the data storage, transmission and processing more difficult[1-3]. In addition, due to the "hugh phenomenon", the result of using all bands to classify objects is not satisfactory. In order to make good use of hyperspectral remote sensing data, a number of scholars have proposed many methods of dimensionality reduction of hyperspectral remote sensing data. In general, these methods can be divided into two categories: feature extraction and band selection[4]. Feature extraction methods include principal component analysis (PCA), singular value decomposition (SVD), minimum noise fraction (MNF), etc[5-7]. The disadvantages of feature extraction are the part of the object's information is lost in the process of transformation and the physical meaning of the transformed data is not clear[8]. Compared with the methods of feature extraction, band selection can retain the original spectral channels physical information, so it has a wide range of applications in many fields[9].

In recent years, heuristic algorithm (HA)[10], especially the meta-heuristic algorithm with natural evolution idea, has raised the upsurge of research again in the world. Since J. Holland proposed the classical genetic algorithm, various novel algorithms such as firefly algorithm, harmony search algorithm and fish swarm algorithm have emerged[11]. In order to solve combinatorial optimization problems, the shuffled frog leaping algorithm (SFLA)[12] was first proposed by Eusuff and Lansey in 2003.

The band selection of hyperspectral remote sensing image is a typical complex high-dimensional optimization problem. Ref.[3] conducted experiments with three supervised and seven unsupervised methods, and a variety of criteria are adopted to guide band selection of hyperspectral data. But all methods fail to take information entropy into account. Ref.[13] used an outer particle swarm to estimate the number of bands selected, and an inner one was used to select the optimal band set. Although that achieves higher classification accuracy, there is some computational redundancy. Artificial physics optimization (APO)[14], invasive weed optimization (IWO)[15], and binary social spiders optimization (BSSO)[16] are used to select the optimal band subset from the original bands of hyperspectral remote sensing images. The methods have some redundancy, and there is still certain promotion space for the classification accuracy.

In this paper, we propose a band selection method based on subspace partition (SP) and particle frog leaping algorithm (PFLA). In order to improve the convergence performance of PFLA, three new evolution strategies are designed to form a probabilistic network extension structure to avoid local convergence. To enhance the classification accuracy, a band selection criterion that integrates information entropy and inter-class separability is proposed. They are considered in the fitness function of the frog simultaneously.

In this paper, the information entropy is measured by

$$H_j = \sum_{I=I_{\min}}^{I_{\max}} p_j(I)\ln p_j(I) , \tag{1}$$

where $p_j(I)$ indicates the probability of the spectral coefficient I appearing in the jth band, and $I_{\min}$ and $I_{\max}$ indicate the minimum and maximum spectral coefficients in the jth band, respectively. And the inter-class

* This work has been supported by the National Natural Science Foundation of China (No.61571149).

** E-mail: liulian0603@126.com.cn

separability is measured by the Jeffries-Matusita distance. The formulas are as follows

$$D_{ij}=\frac{1}{8}(\boldsymbol{\mu}_i-\boldsymbol{\mu}_j)^{\mathrm{T}}\left(\frac{\boldsymbol{\Sigma}_i+\boldsymbol{\Sigma}_j}{2}\right)^{-1}(\boldsymbol{\mu}_i-\boldsymbol{\mu}_j)+\frac{1}{2}\ln\frac{\left|(\boldsymbol{\Sigma}_i+\boldsymbol{\Sigma}_j)/2\right|}{\sqrt{|\boldsymbol{\Sigma}_i|}\times\sqrt{|\boldsymbol{\Sigma}_j|}}, \quad (2)$$

$$J_{ij}=[2\times(1-\mathrm{e}^{-D_{ij}})]^{\frac{1}{2}}, \quad (3)$$

where $\boldsymbol{\mu}_i$ and $\boldsymbol{\mu}_j$ represent the mean vectors for class i and class j of band combination, respectively, and $\boldsymbol{\Sigma}_i$ and $\boldsymbol{\Sigma}_j$ represent the covariance matrices for class i and class j of band combination, respectively.

The criterion of band selection will be designed as follows

$$f=\bar{H}\times\sum_i\sum_j J_{ij}, \quad (4)$$

where $\bar{H}$ represents the average of information entropy of all bands which are selected. So in this paper, the purpose is to find an optimal band subset whose f is the maximum. And then hyperspectral remote sensing image classification is completed with higher precision.

In this paper, the SFLA with novel evolution strategies is used to determine the optimal bands subset. The band selection method of hyperspectral images based on SP and SFLA is designed as follows.

Step 1: Reading data and subspace partitioning. Suppose the size of hyperspectral image data matrix is $M\times N$, and the number of bands is L. The categories of all pixels are known. In order to reduce the correlation between bands, the spectral correlation vectors are referenced as the criterion for subspace partitioning[14]. The commend method of bands selection is selecting a band from each subspace, or selecting several bands in each subspace according to certain proportion.

Step 2: Parameter setting and random initialization of SFLA with novel evolution strategies. The frogs population is assumed to be $P_{\max}$. All frogs are randomly distributed in the D dimensional search space. D is the number of bands selected, which is related to the band selection method based on subspace partitioning. In this paper, SFLA with novel evolution strategies is set to the maximum of the optimization problem:

$$\max f(\boldsymbol{X}_i), i=1,2,...,P_{\max}, \quad (5)$$

where $\boldsymbol{X}_i=[X_{i1},X_{i2},...,X_{iD}]$ denotes the ith frog's position, and $f(\boldsymbol{X}_i)$ stands for the fitness value of the ith frog's position.

Step3: Sorting all frogs according to $f(\boldsymbol{X}_i)$ and dividing sub-swarms. All frogs in the population are arranged in descending order according to $f(\boldsymbol{X}_i)$ (i=1,2,...,$P_{\max}$). Assume all frogs are divided into n sub-swarms, and there are m frogs in each ethnic group. The distribution rules are as follows: the first frog into the first sub-swarm, the second frogs into the second sub-swarms, the nth frog enters the nth sub-swarms, then the (n+1)th frog enters the first sub-swarms, and so on, until all frogs are assigned. In the tth generation, the average position of all frog in the kth sub-swarm is remarked as $\bar{\boldsymbol{X}}_k^t$. The best position of frog is remarked as $\boldsymbol{X}_{bk}^t$ and the worst position of frog is remarked as $\boldsymbol{X}_{wk}^t$ in the kth sub-swarm. And the average position of all $\boldsymbol{X}_{bk}^t$ is remarked as $\bar{\boldsymbol{X}}_b^t$. The best position of all frogs is remarked as $\boldsymbol{X}_g^t$. The design of $f\left(\boldsymbol{X}_i^t\right)$ will directly affect the quality of bands subset selected, and then affect the classification accuracy of the hyperspectral image. To get a better classification effect, in this paper the design of $f\left(\boldsymbol{X}_i^t\right)$ takes two aspects into account: information entropy and inter-class separability. And $f\left(\boldsymbol{X}_i^t\right)$ is calculated by

$$f(X_i^t)=\bar{H}_i^t\times\sum_l\sum_j J_i^t(lj), \quad (6)$$

where $\bar{H}_i^t$ represents the average information entropy of bands subset which is represented by the ith frog's position in the tth generation. $J_i^t(lj)$ represents the Jeffries-Matusita distance between the lth band and the jth band which are represented by the ith frog's position.

Step 4: Updating the worst frog's position in the kth sub-swarm. The original SFLA updates the frogs' positions using three strategies, but there is a weak point that is short in information exchange within sub-swarms. It cannot make full use of the existing information of the current population. So in this paper, the following three novel strategies are designed to update the frogs' position.

In the kth sub-swarm, the frog in the worst position will update its position as follows.

Strategy 1: the frog in the worst position will jump towards the best position $\boldsymbol{X}_{bk}^t$ and the average position of the kth ethnic group $\bar{\boldsymbol{X}}_k^t$ is

$$\boldsymbol{d}_k^t=w_1\times\boldsymbol{r}_1\times(\boldsymbol{X}_{bk}^t-\boldsymbol{X}_{wk}^t)+w_2\times\boldsymbol{r}_2\times(\bar{\boldsymbol{X}}_k^t-\boldsymbol{X}_{wk}^t), \quad (7)$$

$$\boldsymbol{X}_{wk}^{t+1}=\boldsymbol{X}_{wk}^t+\boldsymbol{d}_k^t, \quad (8)$$

where w_1 and w_2 are the weight coefficients, $w_1+w_2=1$. $\boldsymbol{d}_k^t$ is the distance vector of the frog jumping. $\boldsymbol{r}_2=[r_{21},r_{22},...,r_{2D}]$ and $\boldsymbol{r}_1=[r_{11},r_{12},...,r_{1D}]$ are random vectors in D dimensional space. $r_{1j}(j=1,2,...D)$ and $r_{2j}(j=1,2,...D)$ are uniform random numbers between 0 and 1.

Strategy 2: if the fitness value of new position is not better than the previous position's, the frog will jump towards the best position of all frogs $\boldsymbol{X}_g^t$ and the average position of all $\boldsymbol{X}_{bk}^t$ $\bar{\boldsymbol{X}}_b^t$ is as follows

$$\boldsymbol{d}_k^t=w_1\times\boldsymbol{r}_3\times(\boldsymbol{X}_g^t-\boldsymbol{X}_{wk}^t)+w_2\times\boldsymbol{r}_4\times(\bar{\boldsymbol{X}}_b^t-\boldsymbol{X}_{wk}^t), \quad (9)$$

$$\boldsymbol{X}_{wk}^{t+1}=\boldsymbol{X}_{wk}^t+\boldsymbol{d}_k^t, \quad (10)$$

where $\boldsymbol{r}_3=[r_{31},r_{32},...,r_{3D}]$ and $\boldsymbol{r}_4=[r_{41},r_{42},...,r_{4D}]$ are random vectors in D dimensional space. $r_{3j}(j=1,2,...D)$ and $r_{4j}(j=1,2,...D)$ are uniform random numbers between 0 and 1.

Strategy 3: if the fitness value of new position is not better than the original position's, this frog will appear at a random location in the D dimensional search space as

$$\boldsymbol{X}_{wk}^{t+1} = \boldsymbol{r}_5 \times (\boldsymbol{u} - \boldsymbol{l}) + \boldsymbol{l} , \tag{11}$$

where $\boldsymbol{r}_5 = [r_{51}, r_{52}, ..., r_{5D}]$ is a random vector in D dimensional space. $r_{5j}(j=1,2,...D)$ is a uniform random number between 0 and 1. $\boldsymbol{u}$ and $\boldsymbol{l}$ are D dimensional vectors which represent the upper and lower limits of $X_i^t(i=1,2,...,P_{\max})$, respectively.

Step 5: if the iterations do not reach the maximum, go to Step 3, otherwise output the optimal position of frog and its fitness value, and terminate the algorithm.

Hyperspectral AVIRIS data used in the simulation is a hyperspectral image of the agroforestry experimental site in Indiana, northwestern United States, June 1992. In the experiment, 200 bands are used, and the size of hyperspectral image is 144×144. Fig.1 shows the spectral correlation of 200 bands. Refer to this, $\boldsymbol{u} = [37,80,104,145,200]$ and $\boldsymbol{l} = [1,38,81,105,146]$. So in this paper, $D = 5$.

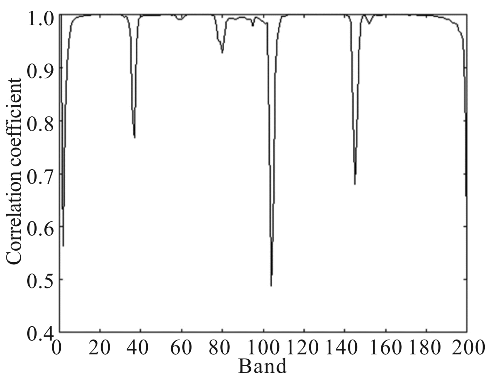


Fig.1 Spectral correlation of 200 bands

In the simulation, the proposed method is compared with the band selection methods based on APO[14], IWO[15], BSSO[16] and SFLA. The common parameters of all algorithms are set as follows. The population size is $P_{\max}$=30, the maximum iteration is $T_{\max}$=100, and the independent trials times are 50. And other main parameters of APO, IWO, BSSO and SFLA are set referring to Refs.[14], [15], [16] and [12], respectively.

In SFLA, n=10, m=3. And the parameters of PFLA are set as follows: n=10, m=3, w_1=2.05, w_2=2.05.

Tab.1 and Tab.2 show the average classification accuracy and average running time of nine objects by using the same algorithm under different fitness functions. The fitness function 1 is proposed in this paper, and the fitness function 2 is put forward in Ref.[11].

Tab.1 Classification accuracy of the same algorithm under different fitness functions

Classification accuracy	APO	IWO	BSSO	SFLA	PFLA
Fitness Function 1	0.731 6	0.728 1	0.678 0	0.732 9	**0.733 8**
Fitness Function 2	0.732 7	0.731 6	0.679 7	0.733 1	**0.735 0**

Tab.2 Running time of the same algorithm under different fitness functions

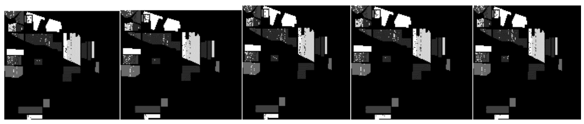
Running time	APO	IWO	BSSO	SFLA	PFLA
Fitness function 1	12.430 1 s	17.966 2 s	14.645 1 s	**6.064 0** s	9.970 9 s
Fitness function 2	6.215 1 s	8.983 1 s	7.322 5 s	**3.032 0** s	4.985 4 s

Through the comparison in Tab.1 and Tab.2, the classification accuracy based on PFLA is the highest in all five algorithms, and the running time of PFLA is shorter than that of APO, IWO and BSSO. And the running time of SFLA is the shortest of all.

The classification results of three, five and nine land objects using different algorithms are shown in Fig.2, Fig.3 and Fig.4, respectively. And Tab.3, Tab.4 and Tab.5 give the classification results under three, five and nine objects in the form of data, where ACA, ART, OBS and OCA represent the average classification accuracy, the average running time, the optimal bands subset and the optimal classification accuracy of 100 independent experiments, respectively.

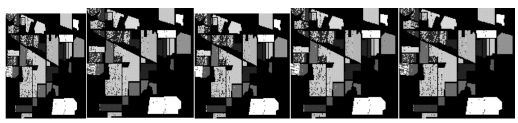
(a) APO (b) IWO (c) BSSO (d) SFLA (e) PFLA

Fig.2 Classification results of three objects using different algorithms



(a) APO (b) IWO (c) BSSO (d) SFLA (e) PFLA

Fig.3 Classification results of five objects using different algorithms



(a) APO (b) IWO (c) BSSO (d) SFLA (e) PFLA

Fig.4 Classification results of nine objects using different algorithms

Tab.3 Classification of three objects using different algorithms

	ACA	*ART*/s	*OBS*	*OCA*
APO	92.54%	17.744 4	[14,42,97,125,162]	93.71%
IWO	92.48%	22.112 4	[14,39,88,119,165]	93.78%
BSSO	86.58%	19.096 9	[7,38,88,131,175]	92.77%
SFLA	92.67%	**7.850 4**	[14,42,89,125,162]	93.64%
PFLA	**92.79%**	11.863 5	[7,44,97,125,165]	**94.00%**

Tab.4 Classification of five objects using different algorithms

	ACA	*ART*/s	*OBS*	*OCA*
APO	82.43%	9.541 3	[34,38,85,131,165]	**84.04%**
IWO	81.20%	12.236 7	[34,42,84,127,162]	83.53%
BSSO	71.29%	11.436 9	[31,38,97,133,161]	82.47%
SFLA	82.23%	**4.618 0**	[24,42,87,127,164]	83.81%
PFLA	**82.80%**	6.841 2	[23,39,87,127,162]	83.99%

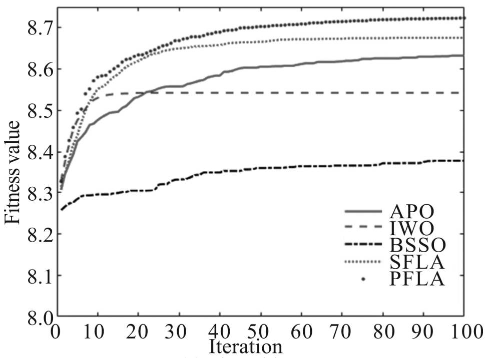
As shown in Tabs.3, 4 and 5, the bands selection method based on SFLA has the best performance in running time, but its performance in classification accuracy is quite the same as that of APO algorithm. On the other hand, the PFLA is slightly inferior to SFLA in running time, but it can achieve higher classification accuracy with shorter time than APO, IWO and BSSO.

In Fig.5, it can be seen that the PFLA is lower in convergence speed compared with SFLA, but it is not easy to fall into the local optimal case and the final convergence precision is the highest. On the other hand, for the five objects, in the 50 independent experiments, the classification precision of optimal bands set found by PFLA is slightly lower than that of the APO algorithm, but its overall average accuracy is the highest.

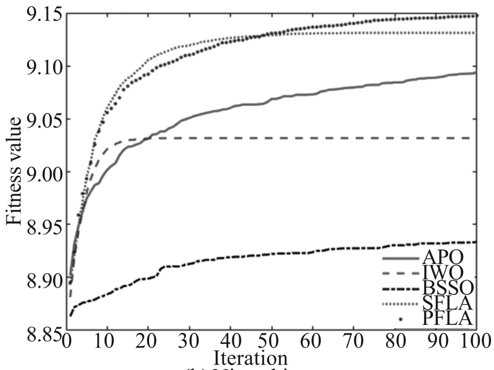
Tab.5 Classification of nine objects using different algorithms

	ACA	*ART*/s	*OBS*	*OCA*
APO	73.27%	25.252 3	[25,39,89,125,162]	74.43%
IWO	73.16%	32.931 3	[34,42,97,135,161]	74.84%
BSSO	67.97%	30.190 3	[23,38,82,125,160]	72.66%
SFLA	73.31%	**11.198 8**	[25,42,89,129,161]	74.35%
PFLA	**73.50%**	17.545 7	[34,38,89,136,161]	**74.95%**

A band selection method for hyperspectral remote sensing images based on subspace partition and particle frog leaping optimization algorithm is proposed in this paper. The simulation results show that the proposed algorithm has certain advantages over the existing similar algorithms in terms of classification accuracy and running time.



(a) Five objects



(b) Nine objects

Fig.5 Fitness curves of five algorithms

References

[1] Feng J, Jiao L C, Zhang X and Sun T, IEEE Transactions on Geoscience & Remote Sensing **52**, 4092 (2014).
[2] Du Q and Yang H, IEEE Geoscience & Remote Sensing Letters **5**, 564 (2008).
[3] Groves P and Bajcsy P, Methodology for Hyperspectral Band and Classification Model Selection, Advances in Techniques for Analysis of Remotely Sensed Data, 2003 IEEE Workshop on, 120 (2003).
[4] Samadzadegan F and Mahmoudi F T, Optimum Band Selection in Hyperspectral Imagery Using Swarm Intelligence Optimization Algorithms, International Conference on Image Information Processing, IEEE, 1 (2011).
[5] Kambhatla N and Leen T K, Neural Computation **9**, 1493 (1997).
[6] Ball J E, Bruce L M and Younan N H, IEEE Geoscience & Remote Sensing Letters **4**, 382 (2007).
[7] Zhu K, Yue L, Yang M, Wang L, Xie B and Cheng Y, Journal of Jilin University **46**, 876 (2016). (in Chinese)
[8] Medjahed S A, Saadi T A, Benyettou A and Ouali M, Egyptian Journal of Remote Sensing & Space Science **19**, 163 (2016).
[9] Bhardwaj K and Patra S, Isprs Journal of Photogrammetry & Remote Sensing **138**, 139 (2018).
[10] Wang M, Wan Y, Ye Z, Gao X and Lai X, A Band Selection Method for Airborne Hyperspectral Image based on Chaotic Binary Coded Gravitational Search Algorithm, Neurocomputing, 2017.
[11] Li Zhang, Kamlesh Mistry, Chee Peng Lim and Siew Chin Neoh, Decision Support Systems, 2017.
[12] Eusuff M M, Lansey K E and Lansey K, Journal of Water Resources Planning & Management **129**, 210 (2003).
[13] Su H, Du Q, Chen G and Du P, IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing **7**, 2659 (2014).
[14] Wang L and Wei F, Journal of Harbin Institute of Technology **45**, 100 (2013).
[15] Pahlavani P, Hasanlou M and Nahr S T, Journal of the Indian Society of Remote Sensing **45**, 1 (2016).
[16] Shukla Urvashi Prakash and Nanda Satyasai Jagannath, Expert Systems With Applications **97**, 336 (2018).