



Science

HYPERSPECTRAL IMAGE CLASSIFICATION BASED ON INTELLIGENT OPTIMIZATION FEATURE SELECTION



Peng Yanbin ^{*1}, Zheng Zhijun ²

^{*1,2} School of Information and Electronic Engineering, Zhejiang University of Science and Technology, China

Abstract

Hyperspectral image classification has always been a hot topic. The problem of "dimension disaster" is caused by the high dimension of pixel points and the lack of labeled training sample points. In order to reduce the data dimension, an intelligent optimization algorithm was proposed for feature selection. The new method introduces the principle of mutual information and symmetric uncertainty, constructs the fitness function, selects the candidate feature set with the intelligent optimization algorithm, and obtains the optimal feature set. The SVM classifier was trained in the optimized feature set. In real hyperspectral data set, the new method was compared with various feature selection methods, and the experimental results showed that the optimal feature set has a high classification accuracy.

Keywords: Hyperspectral Image; Classification; Band Selection; Intelligent Optimization; Mutual Information.

Cite This Article: Peng Yanbin, and Zheng Zhijun. (2020). "HYPER SPECTRAL IMAGE CLASSIFICATION BASED ON INTELLIGENT OPTIMIZATION FEATURE SELECTION." *International Journal of Research - Granthaalayah*, 8(4), 104-111. <https://doi.org/10.29121/granthaalayah.v8.i4.2020.14>.

1. Introduction

Hyperspectral remote sensing technology can obtain a large number of ground object image information at a distance [1-4], and take images of the same surface in different bands. The value of a pixel in different bands forms a spectral curve, according to which we can classify and identify ground objects. On the one hand, a large amount of spectral information is conducive to the classification and identification of ground object targets; on the other hand, due to the high correlation between bands, there are a large number of redundant bands, and a shortage of labeled training samples, which lead to the "dimensional disaster" problem to reduce the accuracy of classification and identification. Therefore, how to carry out dimension reduction and find a feature subset for classification and recognition becomes the key to solve the problem.

In the field of hyperspectral image classification, scholars have proposed a variety of feature (band) selection methods. Rashedul [5] proposed a feature selection method based on segmented principal component analysis to classify hyperspectral images. Yu [6] proposed a dimension specification method based on significant feature extraction. Yang [7] proposed the use of separable non-negative matrix decomposition to achieve hyperspectral band selection. Ding [8] proposed a restricted polymorphic ant colony algorithm for band selection. Xie [9] proposed an unsupervised hyperspectral feature selection method based on fuzzy c-means and gray scale optimizer. Peng [10] proposed a hyperspectral image classification method based on unsupervised feature selection.

The above methods focus on eliminating redundant features one by one and forming feature subsets with the elimination method, without considering the problem from the perspective of global optimization. This paper presents a method of hyperspectral feature selection based on particle swarm optimization. The dimension reduction method introduces the mutual information and symmetric uncertainty principle, constructs the fitness function, then uses the particle swarm optimization algorithm to select the candidate feature set, obtains the optimized feature set and the redundant feature set, and the optimized feature set is used for the subsequent classification and recognition task. In real hyperspectral data set, the new method was compared with various feature selection methods, and the experimental results showed that the optimal feature set has a high classification accuracy.

2. Materials and Methods

The hyperspectral image is a three-dimensional cube, as shown in Fig.1, where I, J and K correspond to their length, width and spectral dimensions respectively. The k th band image is represented as B_k , which is a $I \times J$ dimensional pixel matrix. For one of the pixels, the values of the corresponding points in each band image can be extracted to form an n -dimensional spectral feature. At present, all the bands form the candidate feature set, and the result of feature selection is to obtain the optimal feature subset from which the sample points are formed and the subsequent classification and recognition tasks are completed.

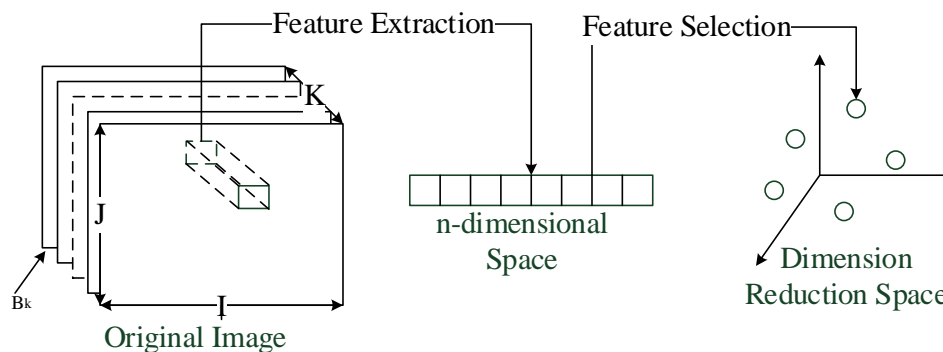


Figure 1: Feature extraction and feature selection

The classification process of hyperspectral images is shown in Fig.2. Firstly, the hyperspectral image is preprocessed to remove the noise band, and the feature set obtained is called candidate feature set B. Then, Particle Swarm Optimization (PSO) [11] was used to select the features from B, and the optimal feature set Q and redundant feature set R were obtained. Finally, the optimal

feature set is used to construct the training set, which is used to train the classifier and classify the new samples.

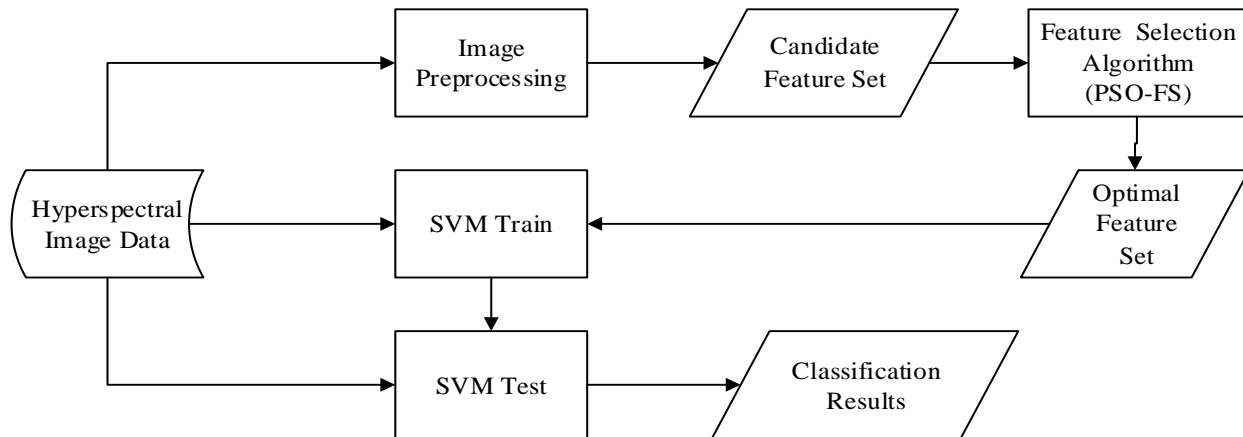


Figure 2: Classification flowchart of hyperspectral image data

The process of band optimization is to remove redundant bands from the candidate feature set and obtain the optimized feature set. Selecting the best feature combination in the candidate feature set can improve the accuracy of hyperspectral classification. The key to feature optimization by using particle swarm optimization algorithm is the particle coding and the selection of fitness function.

Suppose the candidate feature set is $B = \{B_1, B_2, \dots, B_n\}$, the purpose of feature selection is to select a subset $Q = \{Q_1, Q_2, \dots, Q_m\}$ from B, satisfying $Q \subseteq B, m \leq n$. The redundant feature set is $R = B - Q$. A particle is an n-dimensional vector, each dimension of which represents a band. The *i*th dimension component in the particle vector indicates the feature selection result, where 1 means the *i*th feature is selected, and 0 means the feature is not selected. Therefore, the selected band is the band whose component value is 1, as shown in Fig.3. Each value of a particle is the result of a feature selection. For example, in the AVIRIS image, after pre-processing to remove noise bands, there are a total of 200 band images, so a 200-dimensional particle can be used to represent the candidate feature set.

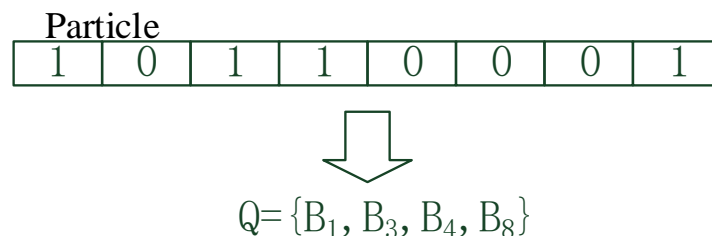


Figure 3: Representation of particle

Feature selection is to find the feature subset with the strongest classification ability in the candidate feature set. Because of the great correlation between the features, the removal of redundant features can improve the algorithm's recognition accuracy and reduce the computational complexity. The particle swarm optimization algorithm is used for selection, so it is necessary to select a fitness function to identify the classification ability of feature subset. In this paper, a

correlation feature evaluation method is adopted, which selects features with low correlation degree to form a feature subset, thus eliminating redundant features. Assuming that the optimal feature subset is Q and the redundant feature subset is R, the fitness function of the particle swarm optimization algorithm is defined as follows:

$$J(Q) = \frac{\sum_{i,j} U(Q_i, R_k)}{\sum_{i,j} U(Q_i, Q_j)}$$

In J, the numerator represents the symmetric uncertainty relationship between optimal features and redundant features, and the denominator represents the symmetric uncertainty relationship between optimal features. U is symmetric uncertainty function. The larger U value indicates that the two feature sets are more correlated, and vice versa. The specific definition is as follows:

$$U(Q_i, Q_j) = \frac{MI(Q_i, Q_j)}{H(Q_i) + H(Q_j)}$$

$$MI(Q_i, Q_j) = H(Q_i) + H(Q_j) - H(Q_i, Q_j)$$

$$H(Q_i) = \sum p(Q_i) \log p(Q_i)^{-1}$$

$$H(Q_i, Q_j) = \sum p(Q_i, Q_j) \log p(Q_i, Q_j)^{-1}$$

Where Q_i and Q_j represents two band images of hyperspectral image data, entropy $H(Q_i)$ and mutual information $MI(Q_i, Q_j)$ are calculated based on the probability distribution of pixel values in the band images, and then the symmetric uncertainty relationship $U(Q_i, Q_j)$ between the two band images is obtained.

After the fitness function J is determined, the particle swarm optimization algorithm (PSO for short) is used to solve the optimization problem. The advantage of PSO is that the algorithm is simple and easy to implement, and the algorithm does not have many parameters to be adjusted. At present, PSO has been widely used in function optimization, neural network training, fuzzy system control and other fields, and achieved good results. In this paper, the optimal feature subset is obtained based on PSO algorithm, which is referred to as PSO-FS. The pseudo-code is as follows:

Algorithm PSO-FS

r=1; //r is the number of cycles

for each particle i

Initialize V_i ; // Initializes the particle velocity vector

Initialize X_i ; // Initializes the particle position vector

pbest $_i$ = X_i // Initializes the locally optimal location

end for

J(gbest) = min{ J(pbest $_i$)}; // Initializes the global optimal location

```

Do{
For i=1 to N
     $V_i = w \times V_i + c_1 \times rand() \times (pbest - X_i) + c_2 \times rand() \times (gbest - X_i)$  ;
     $X_i = update(V_i)$ ;
    If ( $J(X_i) < J(pbest_i)$  )
         $pbest_i = X_i$ ;
End for
 $J(gbest) = \min\{ J(pbest_i)\}$ ;
r++
}while( $r < 500$  and  $|J(gbest_r) - J(gbest_{r-1})| > \xi$  )
Return gbest

```

Where, V_i is the velocity of particle i ; $pbest_i$ is the historical optimal position of particle i ; $gbest$ is the global optimal position; X_i is the position of particle i ; w is the inertia parameter, usually $w = 0.8$; c_1 and c_2 are the learning factor, usually $c_1 = c_2 = 2$; $rand()$ generates a random number between 0 and 1.

The algorithm terminates automatically when one of the two conditions is met: 1) the number of iterations exceeds 500; 2) the minimization criteria are met. The minimization criterion refers to the change of the global optimal fitness produced by two successive iterations is less than the threshold value ξ . The minimization criterion condition is expressed as $|J(gbest_{r+1}) - J(gbest_r)| < \xi$, and $gbest_r$ is the global position generated in the r th iteration. When the condition is satisfied, the algorithm terminates. In this article $\xi = 0.001$.

3. Results and Discussions

In order to verify the effect of the hyperspectral image classification method based on the feature selection of particle swarm optimization (PSO-FS), experiments were carried out on the real hyperspectral data set in this paper. The data set is hyperspectral data of different crops collected by the AVIRIS imaging spectrometer in a field in northwestern Indiana. A total of 220 band images were collected from the data set. Fig.4 is the image of the 25th band of the data set. Each band image size is 145×145 pixels. Among the 220 bands, 20 noise bands (104-108, 150-163, 220) were discarded after pre-processing, and the remaining 200 bands were optimized by particle swarm optimization algorithm. The image contains a total of 16 kinds of land cover information.



Figure 4: AVIRIS dataset

In this section, two experiments were conducted to verify the effectiveness of the new method. The first group of experiments was to compare the optimal features with the candidate features and the redundant features. The second group of experiments compared the optimal features with other feature selection methods.

Experiment 1: Feature selection divides candidate feature set B into two subsets: optimal feature set Q and redundant feature set R. In order to verify the effectiveness of feature optimization results, training samples and test samples of three feature sets were obtained, and SVM was used to classify pixels on three sets. The experimental results are shown in table 1:

Table 1: Accuracy comparison of optimal features, candidate features and redundant features

Classes	Accuracy		
	Optimal feature set Q	Candidate feature set B	Redundant feature set R
Alfalfa	0.8979	0.6324	0.5443
Corn-no till	0.9604	0.8222	0.8576
Corn-min till	0.9600	0.6816	0.7238
Corn	0.8720	0.7205	0.6722
Grass/pasture	0.9308	0.9121	0.9065
Grass/tree	0.9691	0.9584	0.9583
Grass/pasture-mowed	0.9583	0.6913	0.7664
Hay-windrowed	0.9977	0.9886	0.9907
Oats	0.6555	0.5886	0.6224
Soybeans-no till	0.9160	0.7587	0.7774
Soybeans-min till	0.9774	0.8555	0.8358
Soybeans-clean till	0.9818	0.8108	0.7822
Wheat	0.9381	0.8943	0.8844
Woods	0.9243	0.9542	0.9747
Bldg-grass-tree-drives	0.9883	0.7162	0.7079
Stone-steel towers	0.7674	0.5814	0.7091

It can be seen from table 2 that: 1) the classification accuracy of hyperspectral pixel classification with candidate feature set B is not high, because the redundant band improves the dimension of the classification space, and the lack of labeled training samples leads to the “dimensional disaster”, which reduces the identification accuracy. 2) The classification accuracy of the optimal feature set Q is greatly improved compared with that of candidate feature set B and redundant feature set R.

Experiment 2: in order to test the effectiveness of the algorithm in this paper, two relevant band selection algorithms were selected for comparison: Conv-Deconv algorithm [12] and SPCA-nMI algorithm [5].

Table 2: Accuracy comparison of various methods

Classes	Accuracy		
	PSO-FS	Conv-Deconv	SPCA-Nmi
Alfalfa	0.8367	0.7530	0.7142
Corn-no till	0.9659	0.8590	0.8465
Corn-min till	0.9200	0.8146	0.7990
Corn	0.9526	0.7876	0.8161
Grass/pasture	0.9308	0.9017	0.8950
Grass/tree	0.9955	0.9658	0.9538
Grass/pasture- mowed	0.8750	0.7333	0.7833
Hay-windrowed	0.9931	0.9886	0.9931
Oats	0.6666	0.4444	0.5333
Soybeans-no till	0.8956	0.7798	0.8229
Soybeans-min till	0.9954	0.8486	0.8546
Soybeans-clean till	0.9674	0.7848	0.8496
Wheat	0.9947	0.9803	0.9821
Woods	0.9570	0.9364	0.9105
Bldg-grass-tree-drives	0.9208	0.8214	0.8690
Stone-steel towers	0.8470	0.7705	0.7627

Table 2 depicts the classification accuracy comparison of algorithm Conv-Deconv, SPCA-nMI and OPS-FS. Among them, the performance of OPS-FS algorithm is better than other algorithms.

4. Conclusions and Recommendations

In order to improve the accuracy of hyperspectral image classification and recognition, the first problem to be solved is how to effectively describe the feature representation of pixels. The high dimensionality of hyperspectral data makes it necessary to select the appropriate feature subset from the original feature set to reduce the redundancy of feature. In this paper, the selection of hyperspectral features is transformed into a combinatorial optimization problem, and the fitness function is constructed by using the symmetric uncertainty principle and mutual information theory, and then the particle swarm optimization algorithm is used to solve the optimization problem. The experimental results show that the optimized feature improves the accuracy of hyperspectral image classification.

Acknowledgements

Project supported by the National Natural Science Foundation of China (61803337)

References

- [1] Qishuo Gao, Samsung Lim, Xiuping Jia. Spectral-Spatial Hyperspectral Image Classification Using a Multiscale Conservative Smoothing Scheme and Adaptive Sparse Representation. IEEE Transactions on Geoscience and Remote Sensing, 2019.
- [2] Xiangpo Wei, Xuchu Yu, Bing Liu. Convolutional neural networks and local binary patterns for hyperspectral image classification. European Journal of Remote Sensing, 2019, 52(1):448-462.

- [3] Aili Wang, Ying Wang, Yushi Chen. Hyperspectral image classification based on convolutional neural network and random forest. *Remote Sensing Letters*, 2019, 10(11):1086-1094.
- [4] R. Venkatesan, S. Prabu. Hyperspectral Image Features Classification Using Deep Learning Recurrent Neural Networks. *Journal of Medical Systems*, 2019, 43(7).
- [5] Md Rashedul Islam, Boshir Ahmed, Md Ali Hossain. Feature Reduction Based on Segmented Principal Component Analysis for Hyperspectral Images Classification. 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE). IEEE, 2019.
- [6] Xuchu Yu, Ruirui Wang, Bing Liu. Salient feature extraction for hyperspectral image classification. *Remote Sensing Letters*, 2019, 10(6):553-562.
- [7] YANG Gang, SUN Weiwei, ZHANG Dianfa. Separable Nonnegative Matrix Factorization Based Band Selection for Hyperspectral Imagery. *Journal of wuhan university*, 2019, 44(5): 737-744.
- [8] Xiaohui Ding, Shuqing Zhang, Huapeng Li. A restrictive polymorphic ant colony algorithm for the optimal band selection of hyperspectral remote sensing images. *International Journal of Remote Sensing*, 2019(2):1-25.
- [9] Fuding Xie, Cunkuan Lei, Fangfei Li. Unsupervised hyperspectral feature selection based on Fuzzy c-means and Grey Wolf Optimizer. *International Journal of Remote Sensing*, 2019, 40(9):3344–3367.
- [10] Peng yanbin, qiu weiwei, zheng zhijun. A method of hyperspectral images classification based on unsupervised feature selection. *Journal of Optoelectronics Laser*, 2018, 29 (8): 103-108.
- [11] Sheng-Yong Du, Zhao-Guang Liu. Hybridizing Particle Swarm Optimization with JADE for continuous optimization. *Multimedia Tools and Applications*, 2019(3).
- [12] Mou L, Ghamisi P, Zhu X. Unsupervised Spectral-Spatial Feature Learning via Deep Residual Conv-Deconv Network for Hyperspectral Image Classification. *IEEE Transactions on Geoscience & Remote Sensing*, 2018.

*Corresponding author.

E-mail address: pyb2010@126.com