Partition Optimal Band Selection Method for Hyperspectral Image

Yuetao Pan¹, Shishuai Xing¹ and Danfeng Liu¹*

¹ College of Information and Communication Engineering, Dalian Minzu University, Dalian, Liaoning, 116600, China

*Corresponding author’s e-mail: liudanfeng@dlnu.edu.cn

Abstract: Hyperspectral data has the characteristics of large number of bands, narrow band width and large amount of data, which brings great difficulties to the further interpretation of the image. The partition optimal band selection (POBS) is proposed to select a certain number of bands according to the actual needs. The weight calculation formula of the band includes three parameters: standard deviation, information entropy and correlation coefficient, so it can be guaranteed that the selected band subset contains a large amount of information and low correlation. In addition, the whole band space is evenly divided into several band subspaces, and then the band with the largest weight is selected in each band subspace. The bands can be more dispersed through the uniform partition of the band space. It can ensure that the bands in the band subset are more representative. The classification test is carried out on two hyperspectral datasets by KNN method, and the classification results show that POBS selects a large amount of band information, and the divisibility of the ground object is good.

1. Introduction

Remote sensing is a comprehensive technology developed in the sixties of the 20th century to observe ground objects. It is a technology that obtains relevant data from a target through a spectrometer, and then analyzes and processes the information to obtain the required information. Hyperspectral remote sensing uses many narrow electromagnetic bands to obtain relevant information about the target object. Hyperspectral remote sensing images have important practical significance and research value in target classification and recognition. However, the higher spectral resolution of hyperspectral images is at the cost of higher data and larger amount of data. This brings a lot of problems to the processing of hyperspectral images, such as the increase of information redundancy, the large space needed for data storage, the long processing time of data, and the phenomenon of "disaster of dimensionality" when the number of image samples is small. Therefore, in the case of ensuring that the target information is not lost as much as possible, a subset of bands containing most of the structure and information of tens to hundreds of bands is selected to remove some bands that are redundant or cause interference.

In recent years, many unsupervised band selection algorithms have been proposed. These methods are mainly divided into two categories, clustering-based and ranking-based. The clustering-based methods, for example, fast density-peak-based clustering (E-FDPC)[1] and Ward’s linkage strategy using divergence[2], each band is regarded as a point, and the similarity matrix is constructed by Euclidean distance or other criteria. According to the matrix, some clustering methods[3] are used to cluster these bands into a group, so that the bands with high similarity are grouped into a group. The most relevant band with other bands in each cluster is finally selected as representative band [4-5].

The ranking-based method, for example, typical methods have maximum-variance principal
component analysis (MVPCA)[6] and linearly constrained minimum variance based on constrained band selection (LCMV-CBS)[7]. This kind of method is based on some standard to calculate the score of each hyperspectral band. Then the scores are sorted in ascending or descending order, and the band corresponding to the maximum or minimum integral value is selected as the target subset according to the meaning of the score[8].

For hyperspectral data with continuous spectrum, it is necessary to consider not only the amount of information in the overall band and the correlation between bands, but also the divisibility of objects in the band. The continuous spectral characteristics of hyperspectral data provide information for the divisibility of ground objects. The divisibility of each feature in the band is not absolute, among several kinds of features, it is possible that the divisibility of some features in a certain band or band subset is better, while the divisibility of other features in other bands or band subset is poor, so the band selection should be determined according to the specific research objectives.

In this paper, partition optimal band selection (POBS) is used to select the main band from hundreds of bands. The selected band has a large amount of information, low correlation, and strong divisibility between classes. The research of this method is of great significance not only for the dimensionality reduction of hyperspectral images, but also for the further processing of hyperspectral images.

2. PROPOSED METHOD

Chavez, Berlin, Sowers propose a combined band selection method for optimal index factors[9]. Calculate the OIF of the band combination according to the following formula.

\[
OIF = \frac{\sum_{i=1}^{3} SD_i}{\sum_{j=1}^{3}|CC_j|}
\]  

where, \(SD_i\) is the standard deviation of the \(i\)th band, \(|CC_j|\) is the absolute value of the correlation coefficient. The larger the OIF value, the better the band combination. Sort the OIF values to select the best combination of all combinations. The three-bands combination having the largest OIF because it should display the most information with the least amount of duplication.

However, the OIF method has its limitations, because the farther the bands are separated, the smaller the correlation coefficient is. There is no guarantee that the selected band subset is the optimal band. Another reason is that there are too many hyperspectral bands, each possible band combination needs to calculate its OIF value, the calculation is too complicated. Besides, it is not only necessary to consider the amount of information and the correlation all the bands, but also to consider the divisibility of objects in different bands. The continuous spectral characteristics of hyperspectral data provide information for the divisibility of ground features. When classifying, some features may have better divisibility in a certain band, but poor divisibility in other bands. The bands selected by the OIF method are easily concentrated in a certain range of bands, and it can not select frequency bands in a decentralized manner.

In order to solve these problems, POBS is proposed in this paper, The flow chart is shown in Figure 1.

![Figure 1. The overall structure of POBS](image)

First, the hyperspectral image cube is uniformly divided into multiple partitions in space, and then the band with the highest weight in each partition is selected. Ultimately achieve the purpose of band selection. The principle of this method is that the smaller the correlation between the bands, the larger the standard deviation of the bands, the greater the amount of information in the band combination, and the better the band is finally selected. The information amount of the band combination is inversely proportional to the correlation coefficient between the bands, and proportional to the standard deviation and information entropy of the band.
The hyperspectral image cube is evenly divided into multiple partitions in space, and the bands with high similarity in a certain spectral range are grouped into a group, which can fully mine the context information of the whole band space and obtain a more decentralized selection subset. The band with the largest weight in the group is selected as the band subset. By using the idea of sorting-based method, the bands with low correlation and rich information in each group are obtained. This method can make better use of the local distribution characteristics and make the subset more differentiated.

2.1. Band Grouping
The partition point \( G \) is defined as
\[
G(i) = \begin{cases} 
1, & \frac{(B - \text{mod}(B, b)) \times i}{b}, 2 \leq i \geq b - 1 \\
B, & i = b 
\end{cases}
\]
where \( B \) is the total number of bands, \( b \) is the number of selected bands. Concretely, the \( \{A_i\}_{i=1}^{b} \) is defined as follows. The result of this division is that all bands are grouped evenly in space.
\[
A_i = \{a_j\}_{j=G(i)+1}^{G(i+1)}, i = 1, 2, ..., b-1.
\]

2.2. Maximum Weight
The band weight is defined as
\[
w_i = \begin{cases} 
\sigma_i \times H_i, & \text{count}(A) > 5 \\
|r_{i,i-1}| + |r_{i,i+1}|, & \text{otherwise} 
\end{cases}
\]
where \( |r_{i,i-1}| \) is the absolute value of the correlation coefficient between the \( i \)th band and the previous band, \( |r_{i,i+1}| \) is the absolute value of the correlation coefficient between the \( i \)th band and the latter band. \( \sigma_i \) is the standard deviation of the \( i \)th band. \( H_i \) is the information entropy of the \( i \)th band.

The Standard deviation is defined as
\[
\sigma = \sqrt{\frac{1}{M \times N - 1} \sum_{i=1}^{M \times N} (x_i - \overline{x})}
\]

The correlation coefficients is defined as
\[
r_{x,y} = \frac{\sum (x_i - \overline{x})(y_j - \overline{y})}{\sqrt\sum(x_i - \overline{x})^2 \sum(y_j - \overline{y})^2}
\]

Information entropy (IE)[10]. IE is a noise insensitive criterion to measure the information hidden in a stochastic variable. For a band, its information entropy is defined as
\[
H = -\sum_{z} p(z) \log p(z)
\]
where \( \Omega \) represents the subspace of the band space, \( p(z) \) indicates the possibility that event \( z \in \Omega \) appears in the image.

3. Experiment

3.1. Data set
Pavia University Scene. Pavia University Scene was acquired by the Reflective Optics System Imaging Spectrometer (ROSIS) system during a flight campaign over Pavia, Northern Italy in 2002. It consists of 610×340 pixels and 103 bands, 9 classes of land cover objects.

Indian Pines Scene. Indian Pines Scene was captured by AVIRIS sensor in North-Western Indiana in 1992. It consists of 145×145 pixels and 200 spectral reflectance bands in the wavelength range of
0.4−2.5μm. There are 16 classes of objects contained in the image.

3.2. Experimental Setup
The K-Nearest Neighbors (KNN) classifier is used to test the classification accuracy. In the experiment, 10% of each sample is randomly selected to train the classifier, and the rest is used for testing. In order to facilitate verification, this test experiment selects 10 bands from each of the two commonly used hyperspectral data sets as band subsets.
- All bands participate in classification.
- Ten bands are randomly selected from all bands for classification.
- Without partition, the first 10 bands with the largest weight are directly selected for classification.
- According to POBS, ten bands are selected for classification.

3.3. Results
An appropriate number of bands can be selected according to POBS according to requirements. In this experiment, in order to facilitate the verification of the effect and shorten the experiment time, 10 bands are selected for experimental testing.

According to our formula for calculating the weight of each band, calculate the weight of all the bands of the two data sets, so that the band can be selected according to the weight. Figure 2 and 3 show the weight of each band. The abscissa represents each band, and the ordinate represents the weight of each band. In Pavia University Scene, according to POBS, (11, 21, 22, 41, 52, 62, 63, 82, 92, 93) is selected.

In Indian Pines Scene, according to POBS, (22, 23, 45, 88, 89, 132, 133, 155, 177, 199) is selected.

![Figure 2. Weight of each band on Pavia University](image1)

![Figure 3. Weight of each band on Indian Pines Scene dataset](image2)

With the help of KNN method, four different classification strategies are used on two data sets to calculate the accuracy of classification and the time spent.
Table 1. Accuracy and time of classification using KNN

| Strategy            | Pavia University Scene | Indian Pines Scene |
|---------------------|------------------------|--------------------|
|                     | Accuracy   | Time   | Accuracy | Time   |
| All Bands           | 86%        | 380ms  | 67%      | 42ms   |
| Random Selection    | 58%        | 37ms   | 43%      | 6.2ms  |
| Sort Selection      | 67%        | 38ms   | 64%      | 3.9ms  |
| POBS                | 85%        | 37ms   | 66%      | 3.7ms  |

3.4. Result Analysis

Through the comparative experiments of the four strategies on two hyperspectral data sets, it can be seen that the classification accuracy of the bands selected by the zoning optimal band selection method is roughly the same as that of all bands participating in classification, but the time required for all bands to participate in classification is much longer than that of 10 bands. The classification accuracy of the bands selected by the zoning optimal band selection method is higher than that of randomly selecting 10 bands from all bands for classification and directly selecting the first 10 bands with the largest weight for classification without partition. It can be proved that the band selected by the partition optimal band selection method has the advantages of large amount of band information, low correlation, strong divisibility between classes and higher efficiency.

4. CONCLUSION

The continuous spectral characteristics of hyperspectral data provide information for the divisibility of ground features. POBS selects the main band subset from all the bands of the hyperspectral image, which can not only greatly reduce the data dimension of the hyperspectral image, but also retain useful information more completely. This method not only considers the amount of information in the band and the correlation between the bands, but also considers the divisibility of objects in different bands.

POBS fully considers the amount of information contained in the hyperspectral image and the correlation of all the bands, and also optimizes the band selection by partitioning. The KNN classifier is used to classify the 10 bands selected by the POBS method to test its accuracy. The experimental results show that the partition optimal band selection method can select the ideal band, and can improve the effect of image classification, and this method is simple, short operation time and greater effect. So it is a better band selection method. Future works include automatically give the number of recommended bands. Using smarter grouping methods would also deserve some attention.

REFERENCES

[1] S. Jia, G. Tang, J. Zhu, and Q. Li, (2016) “A novel ranking-based clustering approach for hyperspectral band selection,” IEEE Trans. Geosci. Remote Sens., vol. 54, no. 1, pp. 88–102.
[2] A. Martínez-UsóMartinez-Uso, F. Pla, J. M. Sotoca, and P. Garcia-Sevilla, (2007) “Clustering-based hyperspectral band selection using information measures,” IEEE Trans. Geosci. Remote Sens., vol. 45, no. 12, pp. 4158–4171.
[3] C.-I. Chang, Q. Du, T.-L. Sun, and M. L. G. Althouse, (1999) “A joint band prioritization and band-decorrelation approach to band selection for hyperspectral image classification,” IEEE Trans. Geosci. Remote Sens., vol. 37, no. 6, pp. 2631–2641.
[4] Q. Wang, Z. Qin, F. Nie, and X. Li, (2019) “Spectral embedded adaptive neighbors clustering,” IEEE Trans. Neural Netw. Learn. Syst., vol. 30, no. 4, pp. 1265–1271.
[5] Y. Qian, F. Yao, and S. Jia, (2009) “Band selection for hyperspectral imagery using affinity propagation,” IET Comput. Vis., vol. 3, no. 4, pp. 213–222.
[6] A. Martínez-UsóMartinez-Uso, F. Pla, J. M. Sotoca, and P. Garcia-Sevilla, (2006) “Clustering-based multispectral band selection using mutual information,” in Proc. 18th Int. Conf. Pattern Recognit. (ICPR), pp. 760–763.
[7] C.-I. Chang and S. Wang, (2006) “Constrained band selection for hyperspectral imagery,” IEEE Trans. Geosci. Remote Sens., vol. 44, no. 6, pp. 1575–1585.
[8] K. Sun, X. Geng, and L. Ji, (2015) “Exemplar component analysis: A fast band selection method for hyperspectral imagery,” IEEE Geosci. Remote Sens. Lett., vol. 12, no. 5, pp. 998–1002.
[9] Chavez P. S, Berlin G. L, Sowers L. B. (1984) Statistical Method for Selecting Landsat MSS Ratios[J]. Journal of applied photographic engineering, 8(1):23-30.
[10] C. E. Shannon, (1948) A mathematical theory of communication. ACM.