Band Selection of Hyperspectral Images Based on Bhattacharyya Distance

CAI SIMIN, ZHANG RONGQUN, CHENG WENLING, YUAN HUI
Department of Information and Electrical Engineering
China Agricultural University
Tsinghua East Road No. 17, Beijing 100083
The People's Republic of China
caisimin123@163.com
Corresponding author: ZHANG RONGQUN
E-Mail: zhangrq@cau.edu.cn

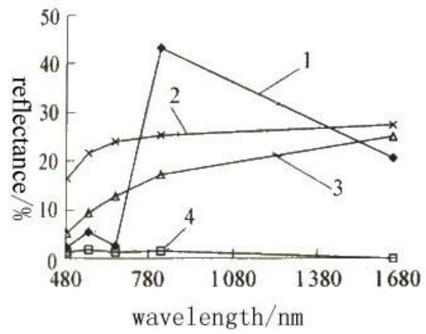
Abstract: With the development of sensor technology, the spectral resolution of remote sensing image is continuously improved. The appearance of the hyperspectral remote sensing is a tremendous leap in the field of remote sensing. The increasing availability of hyperspectral data and image has enriched us with better and finer data and it also enable us a much stronger ability to identify features. However the approaches in the feature identify of hyperspectral images are not as successful as we thought. Too many bands and a large amount of data not only bring difficulties in data storage and transmission, but also bring new challenges in hyperspectral image processing technology, especially the hyperspectral image feature recognize. Band selection aims to recognize the features effectively. We should distinguish the features by utilizing their spectral curve properties. These curves are found to have important information to recognize the different land cover types. So it makes great sense to choose the best combination of many bands and form a new hyperspectral image space. This procedure is usually called features selection. Bhattacharyya-Distance is one of the commonly used methods. It is one kind of statistic distance. It can more reasonably measure the distance between different land types in super multi-dimensional space. The hyperspectral data used in this paper is obtained by the sensors OMIS (Operational Modular Imaging Spectrometer). In this paper, we propose a band selection method based on the Bhattacharyya distance. In the proposed method, we try to find the optimize band combination. We divide land types in the research area into five classes (the five classes are seawater, fishery, building, vegetation and crops); calculate the Bhattacharyya distance between the five class pairs. According the optimal band subset selected by the Bhattacharyya distance, we make a classification and evaluate the classification accuracy. Experimental results show that the proposed band selection method compares favorably with conventional methods.
Key-Words: Hyperspectral images; Band selection; Bhattacharyya distance; Coincident spectral plot

1 Introduction

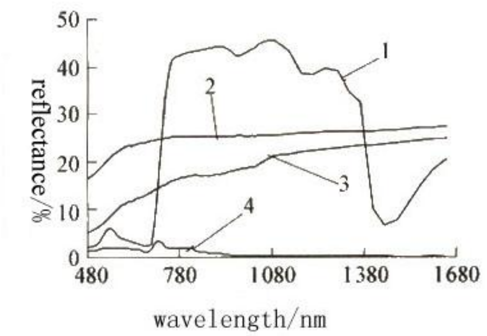
Multispectral sensors have been widely used since the 1960s. The commonly used multispectral images are LandSat TM, ETM+ or ASTER data. However, these images have only a few or dozens of spectrum channels (bands), the spectral resolution is generally around 100 nanometers. That is to say, traditional sensors can only collect spectral data less than twenty bands due to the inadequate sensor technology [1]. In recent years, spectral image sensors have been improved to collect spectral data in several hundreds bands, which are called hyperspectral image scanners. The increasing availability of hyperspectral data and image has enriched us with better data for environmental monitoring, forest tree species, geological exploration and many other applications as well. Currently, most studies focus on either using airborne hyperspectral sensors such as the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). The AVRIS scanners developed by JPL of NASA provide 224 contiguous spectral channels. The AVIRIS was put into use in 1987. It is a typical and maturity acquisition system of hyperspectral data and achieved great success in the short span of two decades.

The development of hyperspectral remote sensing in China synchronizes with other international countries [2]. Shanghai Physics Technical Institute developed a sub-Infrared Spectromter in cooperation with the United States GER in 1983. They developed Operative Modular Airborne Imaging Spectrometer (OMIS) in 1996. The system has two kind of working modular. One is OMIS-Ⅰ, the other is OMIS-Ⅱ. It is a unique optical sensor that delivers calibrated images of the upwelling spectral radiance. OMIS-Ⅰ has 128 contiguous spectral channels (bands) with wavelengths from 460 to 12500 nanometers. And OMIS-Ⅱ has 68 contiguous spectral channels (bands) with wavelengths from 460 to 12500 nanometers. In

the same year, another kind of hyperspectral sensor: Pushbroom Hyperspectral Imaging Spectrometer (PHI) was developed. It had 224 contiguous spectral channels (bands) with wavelengths from 400 to 950 nanometers. The latter two, OMIS and PHI, are now commonly used in many applications in the domestic of China. With hyperspectral scanner onboard satellites, hyperspectral image analysis has already been a major area of research and applications.
Hyperspectral images contain richer and finer spectral information than multispectral images. It has a much stronger ability to identify features than multispectral images (Figure 1).



(a) Multispectral curve



(b) Hyperspectral curve
1 Vegetation 2. Road 3. Bare soil 4. Water body

Figure 1. The comparison of multispectral curve and hyperspectral curve

In Figure 1, we can see clearly that hyperspectral curve is more smooth and delicate than the multispectral curve, and thus the interested land types are easier to be recognized by using hyperspectral data than multispectral data. Even more, hyperspectral data can distinguish the small differences of different features, so we can easily distinguish sub type of vegetation. For example, we can tell the vegetation is grass or trees or other kind of vegetation. From the multispectral and hyperspectral curve, we know that hyperspectral image can refine the subtle difference between different features. Theoretically speaking, using hyperspectral images should increase our abilities in classifying land types. However, the remote sensing image classification approach that has ever been successfully applied to multispectral data in the past is not as effective for hyperspectral images as well.
Hyperspectral data aims to analyze and recognize features not only from spatial aspect, but also from spectral aspects [3]. Hyperspectral data has a much higher spectral resolution and a narrower band ranges of each channel. This is the advantage of hyperspectral data. The neighboring bands have a high relevance; consequently a lot of redundant information is generated [4]. As the number of bands increase, the dimensionality of the feature space increase as well, the number of training samples needed for classification subsequent increase. Due to the characteristics of hyperspectral data, reducing dimensionality without loss of class separability has been an important research issue in hyperspectral data classification. That is to say we should appropriately select spectral bands in order to provide information about the features of interested. Therefore, for hyperspectral remote sensing images, it is necessary to select the optimal bands in the specific application. This can not only reduce the high dimensional data space, but also get access to more reliable and accurate classification results. No doubt this will contribute to the classification and feature recognition [5].

Usually, there are four ways to improve the classification performance: improving the class separability, improving the training sample quality, reducing dimensionality, and choosing a good classifier [6]. To improve classification performance, our attention in the paper is focused on dimensionality reduction.

Dimensionality reduction mainly aims to compress the bands, keeping information as much as possible at the same time. Actually there are two ways to reduce redundancy information [7]. The first approach is to use all the data from original feature space and map the effective features and useful information to a lower-dimensional subspace. This method is referred to as feature extraction. The goal of employing feature extraction is to remove the redundant information substantially without sacrificing significant information. Principal component transform analysis (PCA) [8, 9], projection pursuit [10] and feature extraction based pm separability criteria are three commonly used methods. Feature extraction is based on the reorganization and optimization of each spectral band. It can effectively reduce the amount of data; thereby the volume of computation will be contained. This approach using a mapping process, the dimensional space is reduced, but some information contained in the original bands is lost, which is really unfavorable for processing the hyperspectral images. The other approach is feature selection, also called band selection. It is to select a small subset of features which could contribute to class separability. It does not need to do any transform and maintain the original spectral characteristics.

By selecting the best bands combination, we can reform a new hyperspectral images space without loss any important information, which may well represent the other bands. This method directly chooses the features from the original bands for classification characteristics and thus the physical information can be retained. This process is referred as feature selection or band selection. Band selection mainly includes PCA-based band selection [11, 12], information-entropy-based band selection [13, 14] and class-separability-based band selection [15]. The Bhattacharyya distance is widely used in Pattern Recognition as a criterion for Feature Selection.

While interpreting the hyperspectral images, we face a big problem: which bands or band combinations are used to distinguish feature classes. The Bhattacharya Distance is proposed as a measurement. Its main idea is to measure the statistic distance of samples of different bands [16]. Our attention in the paper is focused on dimensionality reduction by band selection. In our approach, we try to select the optimal band combination. We divide the research area into five classes; calculate the Bhattacharyya distance between different class pairs. According the optimal subset selected by the Bhattacharyya distance, we make a classification and evaluate the classification accuracy. This not only functions effectively between two classes, but also performs well for multiclass problem.

The principle of band selection is: the bigger the separability degree among various classes is and the easier to separate each individual class. Class separability need to be enhanced in order to perform a reliable image classification.

The Bhattacharyya distance is covered in many texts on statistical pattern recognition. Bhattacharyya-Distance measures the scatter degree of two classes. In parametric form it is expressed as follows：

$$B_{ij} = \frac{1}{8}\left[\mu_i - \mu_j\right]^T \left[\frac{\Sigma_i + \Sigma_j}{2}\right]^{-1} \left[\mu_i - \mu_j\right] + \frac{1}{2} Ln \frac{\left|\frac{1}{2}\left[\Sigma_i + \Sigma_j\right]\right|}{\sqrt{\left|\Sigma_i\right|\left|\Sigma_j\right|}} \tag{1}$$

Where μ_i is the mean vector for class i, and Σ_i is the corresponding class covariance matrix. The bigger Bhattacharyya distance of the selected bands combination indicates the better separability of the two classes.

For multi-classes, a commonly used method is to calculate the average Bhattacharyya-Distance between them. That is, calculate the Bhattacharyya distance of all the class pairs and get the average. It is defined as:

$$\overline{B} = \sum_{i=1}^{m}\sum_{j=1}^{m} p(w_i)p(w_j)B_{ij} \tag{2}$$

Where $p(w_i)$ is the weight for class i, and B_{ij} is the Bhattacharyya -Distance for class i and j.

The Bhattacharyya distance is successfully used in engineering and statistical sciences. In the context of control theory and in the study of the problem of signal selection is found superior to the Kullback–Leibler distance [17]. The Bhattacharyya distance is also used in evaluating the features in pattern recognition problem. Furthermore, it has been applied in time series discriminate analysis [18].

In summary, advantages of using the Bhattacharyya distance are that:

- it is computationally very simple;
- And it provides a “smoothed" distance between the two classes in the study, which is more appropriate since we do not believe our data to be truly normally distributed.

The objective of this study is to investigate how many bands and in which spectral ranges these bands are needed to identify the interested land types based on Bhattacharyya distance. The study area is very flat, and this region also has dense water networks. Under this project, we identified the effective bands (the optimal bands combination) to extract the interested land types. The result of supervise classification reveals that the selected bands combination can improve the classification accuracy. In this way, we can suggest suitable bands for feature recognition.

2 Method

2.1 Data and processing system

The study area locates in Yangtze River Delta Region of China. The hyperspectral image used is obtained by Operational Modular Imaging Spectrometer (OMIS-Ⅰ). The OMIS-Ⅰ is developed by Shanghai Institute of Technical Physics, Chinese Academy of Science. It has 128 bands. Its spatial resolution is 4 meters; spectral interval is 10nm and 40 nm. The striped bands are removed because of poor quality. If there is much irrelevant and redundant information present or noisy and unreliable data, then the classification results will be inaccurate [19]. Fourier transform is applied to smoothen the original spectra in order to eliminate noise. Actually there are only 63 bands taking part in the approach (effective bands). The study area relates to a 536 by 1820 pixel image (Figure 2) dealing with the Yangtze River Delta Region feature recognition. the OMIS-Ⅰ image taken on October 2004 and identifies certain geographic features.

The softwares MultiSpec, Erdas and ENVI are used as image data processing systems. And Microsoft offices excel is used as statistic tool.

Table 1. The main technical parameters of OMIS-Ⅰ Imaging Spectrometer

Total bands		128
Spectral range(μm)	Spectral resolution(nm)	bands
0.46~1.1	10	64
1.06~1.70	40	16
2.0~2.5	150	32
3.0~5.0	250	8
8.0~12.5	500	8

2.2 Study area and interested classes

The study area lies in Yangtze River Delta Region of China. The terrain is very flat, river network is dense, and the soil is fertility. In this region, plain takes the majority, about four sixth of the total area, water body takes one sixth, and others are hilly and mountains. This region has mild climate and is abundant with products. And it is well-known as a land of fish and rice. According to the characteristics of this area, and from the angle of urban land type classification, we plan to classify the area as five classes such as seawater, fishery, building, vegetation and crop. The study area is shown in Figure.2.

Preprocessing of satellite remote sensing images is the precondition before us using them. The quality of the preprocessing results has a great impact to the later research results, such as classification accuracy and mapping quality. Before we start the experiment, we first conduct a series of preprocessing to the origin image. These preprocessings are included: atmospheric correction, surface-induced geometric distortions, projection transformation, noise reduction, histogram match and other, etc.

Figure 2. The hyperspectral image of study area

2.3 The experiment

The technology route shows in Figure 3.

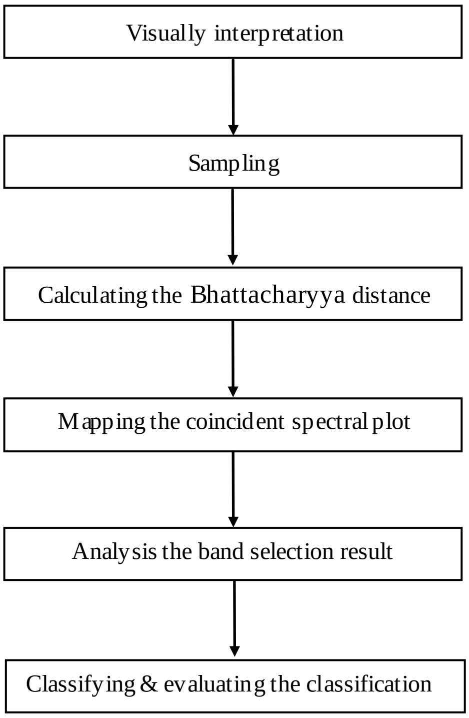


Figure 3. Technology route

Interpretation of remote sensing image is one of an important aspect of remote sensing applications. The applications of hyperspectral remote sensing cannot do without interpretation. Thus we should select the optimal bands and make them into color images which will contain more information. In addition, the band selection of hyperspectral data is not only confined to three bands. In this paper, we select three bands in order to facilitate the display and visual interpretation. The following procedures were adopted:

Step 1 Visual interpreting the image.

Channel 56, 40, 16 are used for color composition display. The land terrain is relatively flat in study area. The land cover of the region is simple and the features distribute regularly. So it is relatively easy to interpret the images visually. In principle the majority of the basic features can be distinguished. The land types are divided into five classes. The five classes are seawater, fishery, building, vegetation and crop. According the result of visual interpreting, we can establish the interpretation signs, which are characteristics to identify feature targets. The interpreter elements used in this experiment are the color, size, shape and location of different features. For example, seawater and fishery are both look blue in the image, but the fisheries look neatly distributed and regularly arranged. That is the difference we can tell the two from each other.

Step 2 Choosing the training samples.

For each class of land types, representative samples are selected for training. Class samples can be very similar to each other in the original spectral measurements and class separability need to be enhanced in order to perform a reliable image classification. The samples should well represent the classes they belonged to. They are selected in the centre of the large target features, avoiding choosing the mixed-pixel. And under each class, 30 training samples are taken. A total of 150 samples are taken for this classification.

Step 3 Calculating the Bhattacharyya distance.

Feature selection is guided by the Bhattacharyya distance measure which is based on a Gaussian distribution model. The reduced dimensionality makes the classification feasible and reliable. We have 5 classes in this approach. It is a multi-classes problem in this experiment (Eq. 2). We should compute the Bhattacharyya distance between each class pair, and count the average value of Bhattacharyya Distance, and then make a sort. According to the principle of the Bhattacharyya Distance, we calculate the best bands combination using the professional software MultiSpec. There are totally 39,711 kinds of band combinations for a group of 3 bands, using the 63 effective bands. The combination (13, 21, and 53) has the maximum average Bhattacharyya distance, so the optimum bands to discriminate classes are given. Part of the result is showed in Table 2.

Table 1 Symbol of different classes

This table shows that the number from 1 to 5 stands for class seawater, fishery, building, vegetation and crop in order.

Channels used: 1-63	
Classes used:	Symbol
1:seawater	1
2:fishery	2
3:building	3
4:vegetation	4
5:crop	5

Table 2 the result of three bands combination
This table gives the top ten band combination of three bands.

		Class pair symbols		12	13	14	15	23	24	25	34	35	45
		Weighting factor		(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)	(10)
	channels	min	ave	Weighted Interclass Distance Measures									
1	13 21 53	1.39	5.43	1.49	4.10	6.21	12.9	4.32	4.96	11.3	2.45	5.09	1.39
2	14 21 54	1.38	5.83	1.69	4.50	6.95	13.7	4.59	5.72	12.2	2.39	5.01	1.38
3	16 22 54	1.37	5.44	1.38	3.91	6.47	13.3	3.58	4.51	11.7	2.64	5.53	1.37
4	16 21 54	1.36	5.35	1.37	3.71	6.47	13.0	3.37	4.30	11.6	2.60	5.57	1.36
5	17 21 53	1.36	5.98	1.64	5.09	6.91	13.7	5.39	5.74	12.0	2.55	5.35	1.36
6	19 21 53	1.36	5.96	1.66	4.96	6.95	13.7	5.12	5.59	11.9	2.65	5.64	1.36
7	18 21 53	1.36	6.05	1.65	5.28	6.93	13.7	5.72	5.77	12.0	2.60	5.44	1.36
8	16 21 53	1.36	5.79	1.68	4.42	7.02	13.7	4.46	5.38	11.9	2.55	5.34	1.36
9	15 21 53	1.35	5.30	1.35	3.92	6.30	12.8	3.77	4.39	11.2	2.50	5.24	1.41
10	15 21 53	1.35	5.82	1.79	4.30	7.09	13.8	4.50	5.52	12.0	2.53	5.29	1.35

Step 4 Map the coincident spectral plot.

First, we obtain the statistic data of each class and each channel. These statistic features include mean value and variance. By Erdas, we can obtain the statistic data of each sample. These data include the minimum value, maximum value, average value, standard deviation spectral value. Using the statistic tools in Erdas to get the statistic data of each sample, we can output these data to Microsoft Excel, to calculate the average statistic data of each class. With these data we can map the coincident spectral plot. Then display the results in the chart, which is the coincident spectral plot. See Figure 4.

Coincident spectral plot, also called multi-bands response chart. This chart is built on the basis of analysis of statistic data. In this plot, we draw the average spectral response for each class in each band. Different letters stand for different class. The length of line represents the size of the standard deviation. The longer the line of one class is, the more easily to distinguish the class from others. The coincident spectral plot shows vividly the location, distribution range, the degree of discrete as well as separability of different classes in every band. It is a direct simple and effective method to analyze the spectral properties of different classes in quantitative way, select the optimum band combination and establish classification tree effectively in this way [20].

In our experiment, we map and classify the land types of the coincident spectral plot in this region. According to the spectral characteristics of different types, we make use of composite spectral plot for band selection in assistance. Seeing from the following plot, we will get the value of the overlap between the classes, the size of separability in order to determine the establishment of clear boundaries.

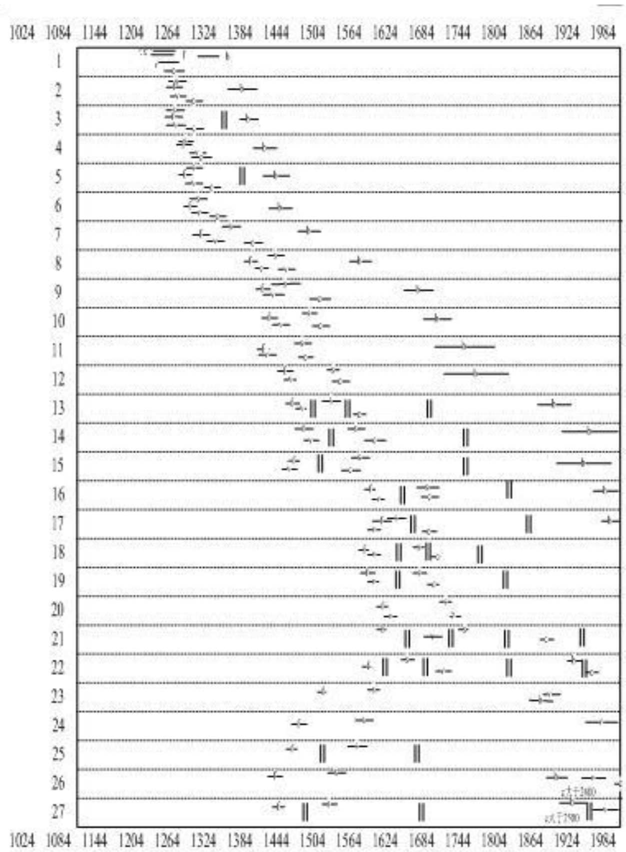


Figure 4. Part of the coincident spectral plot

The remaining the procedures, such as the analysis the band selection result, classifying and evaluating the classification will be detailed in the next part. And the most important point, we shall do a proper evaluation after the experiment [21].

3 Experiment results analysis

3.1 The analysis of all bands standard deviation curve

The standard deviation is used to measure the amount of information, the great the value is, the more information it has, the pixels are easier to distinguish. Great value of standard deviation also means the differences between classes intend to be the largest, and the amount of information is the most abundant. By calculating the standard deviation, combined with the intuitive judge the quality of the image, it is easy to determine which are effective bands and invalid bands.

The distribution of all bands standard deviation curve indicates four peaks mainly concentrated in the following bands ranges: 15-19, 20-25, 26-30, 35-38, 50-55 (Figure 5). The bands combination 13, 21 and 53 selected by Bhattacharyya distance in the previous section generally fall into the range of peaks.

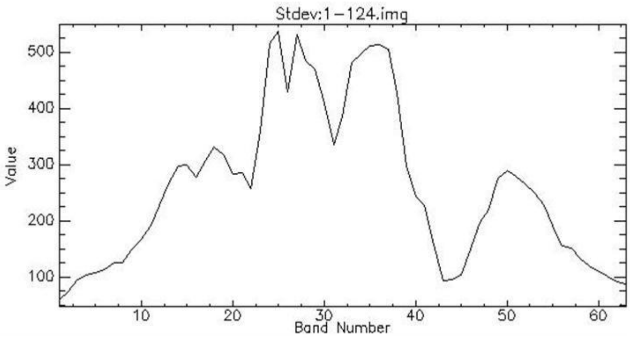


Figure 5. Standard deviation curve for all bands

3.2 Histograms analysis of the selected bands combination

(Note that: in this part, s presents for class seawater, f presents for class fishery, b presents for class building, v presents for class vegetation, c presents for class crop.)

In the following histograms, different colors stand for different classes. Red stands for s class, blue stands for b class, cyan stands for v class, magenta stands for c class and green stands for f class. The horizontal axis shows the data value, the vertical axis represents the number of pixels. In Figure 6 to 8, show the histogram distribution of band 13, 21 and 53.

In Figure 6, channel 13, the pixel number of b class is not particularly large, but the brightness values are a bit large, the majority of which is large than 1800. So it is easy to distinguish the b class from the other four classes (class s, f, v, and c). From the histogram distribution of the channel 53 (Figure 8), it shows clearly that the distribution of class c and v is broad, basically normally. So it is easy to distinguish the two classes according to the data value.

The distribution of class f in channel 21 (Figure 7) only has a narrow range and a small number of pixels, while class s takes up a large quantity. This helps to separate class s from class f. Now the five classes are separated from each other.

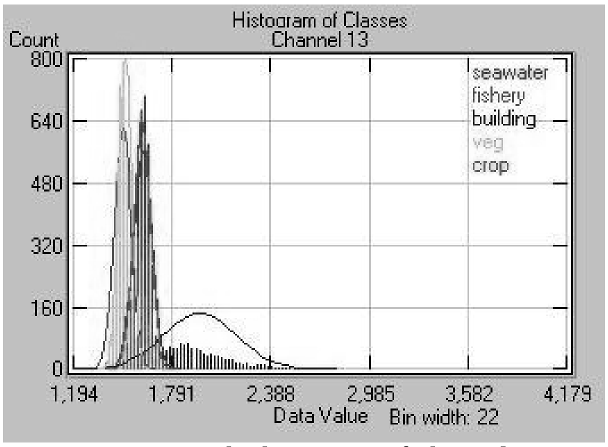


Figure 6. The histogram of channel 13

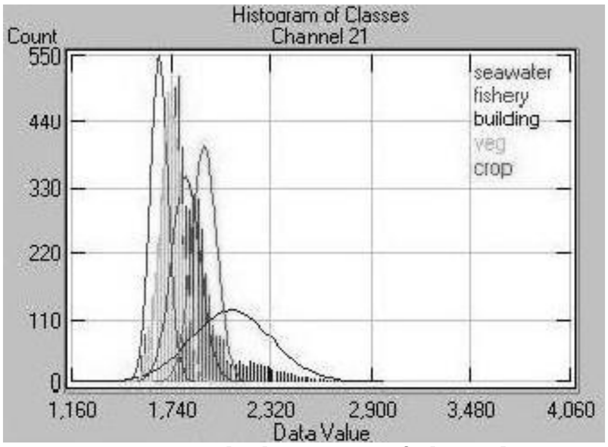


Figure 7. The histogram of channel 21

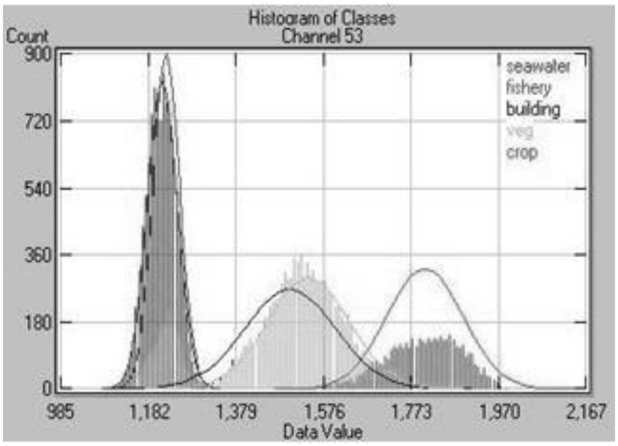


Figure 8. The histogram of channel 53

3.3 Mapping a decision tree according to the coincident spectral plot and analysis

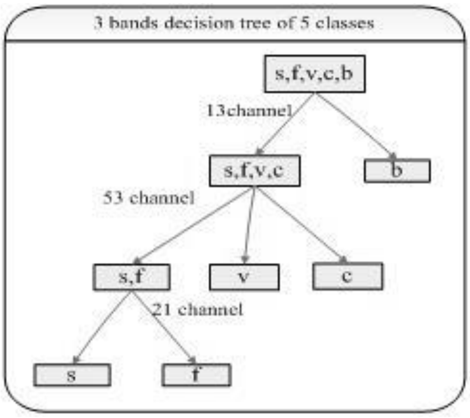


Figure 9. Three bands decision tree of the five classes

From the coincident spectral plot, we can see clearly the distribution characteristics of each band. As for the 3 bands combination (band 13, 21and 53), the five classes can be divided into two categories preliminary by channel 13. Class s, f, v and c form a big class, class b for another. Channel 53 can easily separate class c and v from the big class, which includes class s, f. Then only two classes remaining to separate, that is class s and f. The two are quite similar in spectral properties, and overlap largely in channel 13 and 53, which will take trouble to tell them apart. However, channel 21 can be a very clear distinction between class s and f.

3.4 The results of Classification and accuracy assessment

Several different combinations of bands are selected to test the discriminating power of different land type features. By band selection, we know that the bands 13 21 53 are optimal bands for classification; by classification, the three bands using likelihood relaxation classifier, we obtain a classification map. In order to evaluate the performance of the proposed method, experiment is conducted using all bands to make a comparison, and we also do an classification precision analysis. The classification accuracy result is shown in Table 3, and the classification maps are shown in Figure 10 and Figure 11.

Using the selected optimal 3-band combination generated an average overall accuracy of 68.8%. Using all bands to carry out the classification we will get an average overall accuracy of 65.4%. From the accuracy table we can know that the classification precision of the method in this paper is 3.4% percent higher than all bands maximum-likelihood classifier.

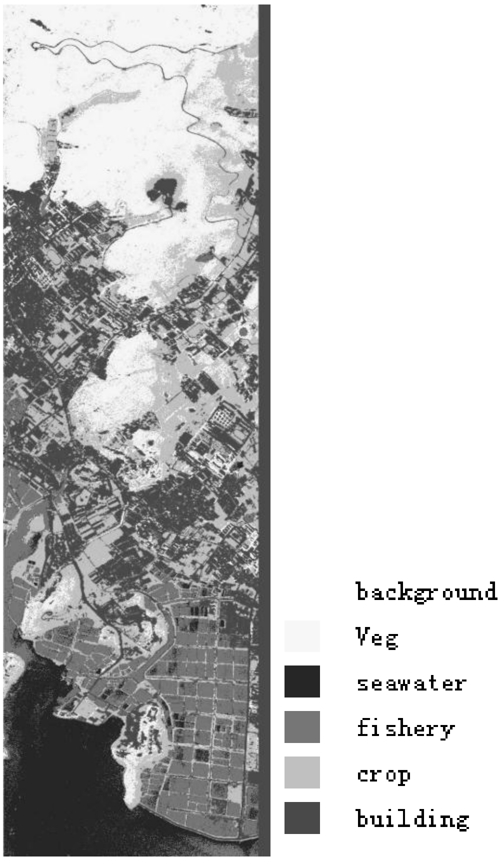


Figure 10. Classification map of three bands

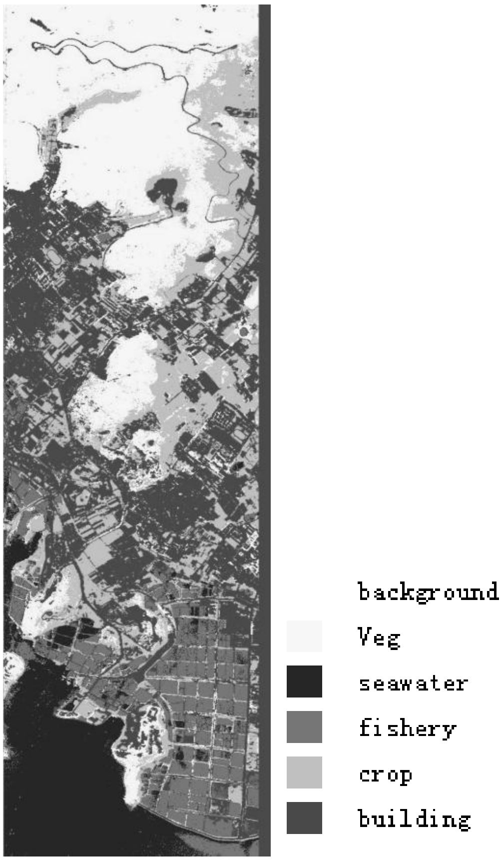


Figure 11. Classification map of all bands

Table 3. Classification accuracy comparisons (units: percent)

Classes	three bands	all bands
Seawater	71.6	74.4
Fishery	77.9	76.2
Building	64.6	63.3
Veg	61.6	60.7
Crop	64.3	63.5
overall accuracy	68.8	65.4

Three bands combination reduce the amount of bands for classification in one hand. In the other hand it can expand the differences in spectral for various classes, and make it much easier to recognize and classify the features. Although for each single class the classification accuracy of three band combination is not increased, for some class there is even a bit of decrease, the overall accuracy is improved. That is true that we use the most effective bands to substitute all bands combination and a shorter computing time to get a better classification result.

4 Conclusion

This study shows that band selection helps improve the classification accuracies for land type recognition. The test data sets (see Figureure 2.) which were delivered by the Operational Modular Imaging Spectrometer (OMIS) in this paper. In feature selection, we used the Bhattacharyya distance method to select the optimal feature combinations.The approach used for classification is the maximum Likelihood Classifier after feature extraction. The number of training samples used in the classification is 30 for each class. For comparison, we use the optimal three bands and all bands separately to do the classification. The optimal three bands classification map is shown in Figure 10. And all bands classification map is shown in Figure 11. From the point of classification accuracy, the former is superior to the latter. The result shows that the band selection based on Bhattacharyya distance is effect to land type classification. The proposed method can not only perform well of two class problem, but also do well with the multiclass problems. This method can exactly reduce the dimensionality of hyperspectral data and increase the accuracy of classification.

Generally speaking, from the aspect of the properties of hyperspectral data, such as large number of bands, large amount of information and high redundancy, we discussed the processing methods. We have provided a method to select the optimum band combination for hyperspectral images. The conclusions of the study are as follows:

a) Hyperspectral remote sensing has become one of the hot areas of research. Hyperspectral data have a high spectral resolution, and thus it is easy to distinguish the objects similar in spectral. However there is a serious problem to face. It is quite difficulty to process the massive hyperspectral data. Here we introduce the concept of feature selection. The goal of employing feature extraction is to reduce the number of features substantially without sacrificing significant information. Thus the accuracy of classification could be preserved and the speed of computation could be reduced.
   Band selection based on Bhattacharyya distance gives a solution. This method is a common practice. Experiment in this paper shows that it is an efficient and convenient method. The result is reliable and validate. It takes both information entropy and the separability with classes into account.
b) The coincident spectral plot is mapped based on the spectral statistic data. This plot shows the location of each class vividly, the distribution range, the degree of discrete as well as separability of different classes in every band. It is a direct simple and effective method to analyze the spectral properties of different classes in quantitative way. On the one hand, the spectral coincident plot can give a visual representation of the useful band or band combination. Decision tree for classification can also be designed to choose different bands or bands combination according to the different needs. On the other hand, some classes, such as corn and soybean, cannot tell from each other due to the large amount of overlap of spectral response prosperities in one single band. In visible spectrum region, class A is greater than class B, and in near infrared region class A is less than class B. This is called reversal phenomenon in spectral response. Class A and class B has a part of overlap in one band, and the overlap is easy to be distinguished in another band. So we can make use of this reversal phenomenon to recognize the features that cannot be tell by only one band. They can be distinguished combining with two or three bands. The coincident spectral plot can intuitively help us to carry out the feature selection of hyperspectral data.
c) The Bhattacharyya distance is successfully used in many fields and statistical sciences. In this paper, the method of band selection of

hyperspectral data based on Bhattacharyya distance is not only a good way to distinguish two classes, but also perform effectively for multi-classes. In allusion to the feature selection of multiclass problems, the average Bhattacharyya distance is proposed to select the band combination. In the new method, the interclass separability is described by this average distance defined by Bhattacharyya distance. The maximum average Bhattacharyya distance gives the optimal band to distinguish the different class pairs. This proposed method will be very helpful to improve the classification accuracy.

d) In this paper, the Bhattacharyya distance to select the optimal bands combination is really a good way to select features. This could not only reduce hte number of features substantially, but also improve the accuracy of classification, as well as save the time of computation while classification.But there is still a problem. The classification accuracy is not significantly increased by this meathod. Therefore the method of feature selection based on the Bhattacharyya Distance needs to be improved.

Acknowledgement

This work was supported in part by the National Natural Science Fund, China (No.40871103). Thanks very much for the help and support is given by Associate Professor Wang Xuhong, the department of Urban and the environment planning, Northwest University 710127, XI' AN, The People's Republic of China.

References:

[1] Cristina Solares, Ana Maria Sanz, Bayesian Networks in the Classification of Multispectral and Hyperspectral Remote Sensing Images, *3rd WSEAS International Conference on REMOTE SENSING*, Venice, Italy, November 21-23, 2007, pp 83-86.

[2] Tong Qingxi, Zheng Lanfen, Xue Yongqi, Zhang Bing, Zhao Yongchao, Liu Liangyun. Hyperspectral remote sensing in China, *Proc. SPIE, 2001, Vol.* 4548, *1*.

[3] He Ting, Wang Jing, Cheng Ye. Study on Geometric Correction of OMIS Images, *Geography and Geo-Information Science*, Vol. 21, 2005, pp. 23-26.

[4] Clark A. Spectral identification by singular value decomposition, *International Journal of Remote Sensing*, vol. 19, 1998, pp: 2317-2329.

[5] N.Kambhatla, R.A.Leen. Dimension reduction by local principal component analysis, *Neural Computation*, vol. 17, 1998, pp: 1493-1516.

[6] Hsien P.F., D. Landgrebe. *Classification of High Dimensional Data*. Ph. D. dissertation, School of Electrical and Computer Engineering, Purdue University, West Lafayette, Indiana,1998.

[7] Weijun Wu, Qigang Gao, Muhong Wang, An Efficient Feature Selection Method for Classification Data Mining, *WSEAS TRANSACTIONSION INFORMATION SCIENCE & APPLICATIONS*, Vol 3, October 2006, pp:2034-2035

[8] D. Manolakardis, D. Marden. Dimension reduction of hyperpectral imaging data using local principal components transforms, *Proceedings of SPIE*. 2004, pp: 393-401.

[9] D.F.Michael, R.M.Mersereau. On the impact of PCA dimension reduction for hyperspectral detection of difficult targets, IEEE *Geoscience and Remote Sensing Letters*, vol. 2, 2005, pp. 192-195.

[10] Jimenez L O., Landgrebe D A. Hyperspectral data analysis and supervised feature reduction via projection pursuit, *IEEE Geoscience and Remote Sensing*, Vol. 37, 1999, pp. 2653-2667.

[11] Pai Hui Hsu. Feature extraction for hyperspectral image, *Proceeding of the 20th Asian Conference Remote Sensing*, 1999, pp. 408-410.

[12] Yang Zhusheng, Guo Lei, Luo Xin. Research on segmented PCA based on band selection algorithm of hyperspectral image, *Engineering of Surveying and Mapping*, vol. 15, 2006, pp. 15-18.

[13] Chavez P S, Berlin G L.Sowers L B., Statistical Method for Selecting Landsat MSS Ratios, *Journal of Applied Photogrammetric Engineering*, vol. 1, 1982, pp. 23-30.

[14] Charles Sheffield. Selecting Band Combination from Multispectral Data, *Photogrammetric Engineering and Remote Sensing*, vol. 51, 1985, pp: 681-687.

[15] Jiang Xiaoguang, Lingli,Wang., Changyao. Spectral Characteristics and Feature Selection of Hyperspectral Remote Sensing Data, *Remote Sensing Remote Sensing Technology and Application*, vol. 17, 2002, pp. 59-65

[16] Tong Qingxi, Zhang Bin, Zheng Fenlan. *Hyperspectral Remote Sensing - Principle, Technology and Application*, Beijing: High education Press, 2002, pp. 67-72.

[17]T. Kailath, The divergence and Bhattacharyya distance measures in signal selection, *IEEE Trans. Comm. Techno.*, COM–15, 1967. pp. 52–60

[18]G. Chaudhuri, J.D. Borwankar, P.R.K. Rao, Bhattacharyya distance-based linear discriminant function for stationary time series. *Comm. Statist. (Theory and Methods)*, 1991, pp. 2195–2205

[19]Qian, Shen-En, Lévesque, Josée, Neville, Robert A. Effect of removing random noise of radiance data using smoothing on data compression onboard a hyperspectral satellite. *WSEAS Transactions on Systems*, 2006, vol. 5, n1, pp 219-224.

[20] Zhao Y S. *Remote Sensing Application Analysis Principle and Method*. Beijing：Science Press，2006，pp：232-235

[21]Michael Doumpos, Athina Salappa, Feature Selection Algorithms in Classification Problems: An Experimental Evaluation, *WSEAS TRANSACTIONS on INFORMATION SCIENCE & APPLICATIONS*, Vol 2, February 2005,pp，77-80