Spectral Features Extraction in Hyperspectral RS Data and Its Application to Information Processing*

Du Peijun[1,2], Fang Tao[2], Tang Hong[2], Chen Yongye[2]
1 *Department of RS and GISci, China University of Mining and Technology, Xuzhou City, Jiangsu Province* 221008
2 *Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai* 200030

Abstract Oriented to the demands of hyperspectral RS information processing and applications, spectral features in hyperspectral RS image can be categorized into three scales: point scale, block scale and volume scale. Based on the properties and algorithms of different features, it is proposed that point scale features can be divided into three levels: spectral curve features, spectral transformation features and spectral similarity measure features. Spectral curve features include direct spectra encoding, reflection and absorption features. Spectral transformation features include Normalized Difference of Vegetation Index (NDVI), derivate spectra and other spectral computation features. Spectral similarity measure features include spectral angle (SA), Spectral Information Divergence (SID), spectral distance, correlation coefficient and so on. Based on analysis to those algorithms, several problems about feature extraction, matching and application are discussed further, and it proved that quaternary encoding, spectral angle and SID can be used to information processing effectively.

Keywords Hyperspectral Remote Sensing; Spectral feature; Feature extraction; Information processing
CLCN TP75 **Document Code** A

0 Introduction

Hyperspectral Remote Sensing was one of the most important breakthroughs of Earth Observation System (EOS) in 1990s. It overcomes the limitations of conventional aerial and multi-spectral RS such as less band amount, wide band scope and rough spectral information expression, and can provide RS information with narrow band width, more band amount and fine spectral information, also it can distinguish and identify ground objects from spectral space, so hyperspectral RS has got wide applications in resources, environment, city and ecological fields. Because hyperspectral RS is different from conventional RS information obviously in both information acquisition and information processing, there are many problems should be solved in practice. One of the most important problems is about spectral features extraction and application in hyperspectral RS data including hyperspectral RS image and standard spectral database. Nowadays, studies on hyperspectral are mainly focused on band selection and dimensionality reduction, image classification, mixed pixel decomposition and others, and studies on spectral features are few. In this paper, spectral features extraction and application will be taken as our central topic in order to provide some useful advices to hyperspectral RS applications.

1 Framework of spectral features in hyperspectral RS data

In general, hyperspectral RS image can be expressed by a spatial-spectral data cube (Fig. 1). In this data cube, every coverage expressed the image of one band, and each pixel forms a spectral vector composed of albedo of ground object on every band in spectral dimension, and that vector can be visualized by spectral curve (Fig. 2). Many features can be extracted from spectral vector or curve, and spectral features are the key and basis of hyperspectral RS applications. Also each spectral curve in spectral database can be analyzed with same method. Although there are some algorithms to compute spectral features,

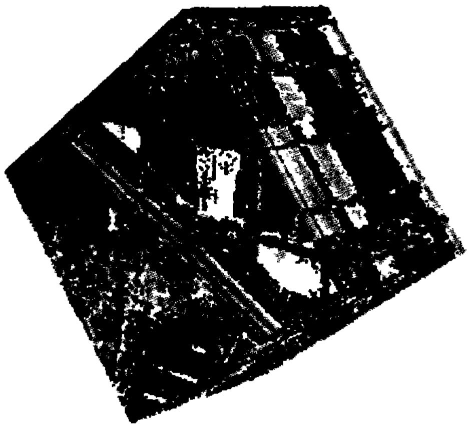

Fig. 1 Hyperspectral image data cube

*The project supported by National High Technology Research and Development Program of China (2001AA135091), National Natural Science Foundation (40401038), China Postdoctoral Science Foundation (2002032152) and Open Research Fund Program of the Key Laboratory of Geomatics and Digital Technology, Shangdong Province (No 2003-03)
Emai: dupj@vip.163.com
Received date: 2003-11-17

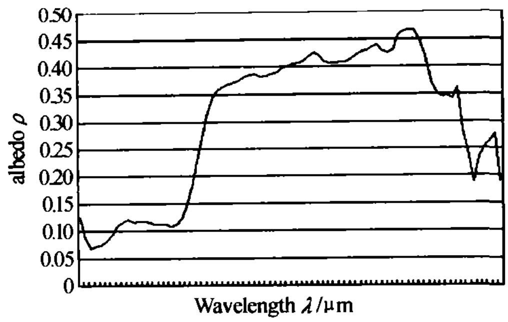

Fig. 2 Reflectance spectral curve of a pixel

the framework and system is still not obvious, so we would like to propose a framework for spectral features in hyperspectral RS data including hyperspectral RS image and standard spectral database.

1.1 Three scales of spectral features

According to the operational objects of extraction algorithms, spectral features can be categorized into three scales: point-scale, block-scale and volume-scale.

Point scale takes pixel and its spectral curve as operational object and some useful features can be extracted from this spectral vector (or spectral curve). In general, hyperspectral RS image takes spectral vector of each pixel as processing object.

Block scale is oriented image block or region. Block is the set of some pixels, and it can be homogeneous or heterogeneous. Homogeneous regions are got by image segmentation and pixels in this region are similar in some given features; heterogeneous region are those image blocks with regular or irregular size, and they are cut from original image directly, for example, an image can be segmented according to quadtree method. In hyperspectral RS image, block scale features can be computed from two aspects. One is to compute texture feature of a block on some characterized bands, and the other is to compute spectral feature of a block. If the block is homogeneous its mean vector can be computed firstly and then spectral of this mean vector can be extracted to describe the block. If the block is heterogeneous, it can be segmented to some homogeneous blocks.

Volume scale combines spatial and spectral features in a whole and extracts features in 3D (row, column and spectra) space. Here, some 3D operational algorithms are needed, for example, 3D wavelet transformation and high order Artificial Neural Network (ANN). Because this type of features is difficult to compute and analyze, we don't research it in current studies.

In this paper, we would like to focus on point scale feature, or those features extracted from spectral vector that may be spectral vector of a pixel or mean vector of a block.

1.2 Three levels of point scale features

From operation object, algorithm principles, feature properties, application modes and other aspects, we think it is feasible to categorize spectral features into three levels: spectral curve features, spectral transformation features and spectral similarity measure features. They are corresponding to analysis on spectral curve with all bands, data transformation and combination with part of all bands and similarity measure of spectral vectors. In our study, data from OMIS and PHI hyperspectral image, USGS spectral database and typical spectra data in China is experimented and two examples are given in this paper. One is to select three regions from PHI image (Region I is vegetation, Region II is built-up land, and Region III is mixed region of some land covers), and the other is spectral curve of three ground objects from USGS spectral database, among them S1 is Actinolite_HS22.3B, S2 is Actinolite_HS116.3B and S3 is Albite_HS66.3B, so S1 and S2 are similar and they are different from S3.

2 Spectral curve features

Spectral curve features are computed by some algorithms based on the spectral curve of certain pixel or ground object, and it can describe shape and properties of the curve. The main methods include direct encoding and feature band analysis.

2.1 Direct encoding

The important idea of spectral curve feature is to emphasize spectral curve shape, so direct encoding is a very convenient method, and binary encoding is used more widely. Its principle is to compare the attribute value at each band of a pixel with a threshold and assign the code of "0" or "1" according to its value. That can be expressed by

$$S[i]=\begin{cases}1 & \text{if } (X_i \geq T)\\ 0 & \text{else}\end{cases}$$

Here, $S[i]$ is code of the ith band, X_i is the original attribute value of this band, and T is the threshold. Generally, threshold is the mean of spectral vector, and it can also be selected by manual method according to curve shape, sometimes median of spectral vector is probably used.

Only one threshold is used in binary encoding, so the divided internal is large and precision is low. In order to improve the approximaty and precision, the quaternary encoding strategy is proposed in this paper. Its primary idea is as follows: (1) the mean of the total pixel spectral vector is computed and denoted by T_0, and the attribute is divided into two internal including $[X_{min}, T_0]$ and $(T_0, X_{max}]$; (2) the pixels located in the two internals are determined and the mean of each internal is got and donated by T_L and T_R, so four internals are formed including $[X_{min}, T_L]$,

$(T_L, T_0]$, $(T_0, T_R]$ and $(T_R, X_{max}]$; (3) each band is assigned one of the code sets {0, 1, 2, 3} according to the internal it is located; (4) to compute the ratio of matched bands number to the total band number as final matching ratio. It proved that quaternary encoding could describe the curve shape more precisely.

Table 1 is the results of analysis to three regions from hyperspectral image, and Table 2 is the results of analysis to data from USGS spectral database. It should be pointed that the results of data from hyperspectral RS image are the mean of all indexes between two pixels that one is from one region and the other is from corresponding region in tables, and the results of data from USGS spectral database are the index value of two spectral curves.

Table 1 Analysis on direct coding to data from hyperspectral image

	Binary encoding			Quaternary encodingI		
	I	Ⅱ	Ⅲ	I	Ⅱ	Ⅲ
I	0.9509	0.9186	0.8667	0.8350	0.6197	0.5102
Ⅱ		0.9642	0.8317		0.8064	0.3810
Ⅲ			0.9701			0.8329

Table 2 Analysis on direct coding to data from USGS spectral database

	Binary encoding			Quaternary encodingI		
	S1	S2	S3	S1	S2	S3
S1	1	0.99	0.52	1	0.89	0.31
S2		1	0.52		1	0.32
S3			1			1

From Table 1 and Table 2 we can know that when binary encoding is adopted, although the ratio in the same region is higher than that between different regions, the difference is not very large. If quarternary encoding is used, the ratio of the same region is smaller than binary encoding, but the ratio between different regions decreased dramatically. So quarternary encoding is more effective in measuring the similarity between different pixels.

Because direct encoding will disperse the continuous albedo into discrete code, the encoding result is affected by threshold obviously and will lead to information loss. Although its operation is very simple, it is only used to some applications requiring low precision, and the threshold should be selected according to different conditions.

2.2 Spectral absorption or reflection feature

Differing from direct encoding in which all bands are used, spectral absorption or reflection feature only emphasizes those bands where valleys or apexes are located. That means those bands with local maximum or minimum in spectral curve should be determined at first and then further analysis can be done. In general, albedo is used to describe the attribute of a pixel, so those bands with local maximum are reflection apex and those with local minimum are absorption valley. The algorithm used to determine apex and valley is as follows

```
for(j = 1;j < N - 1;j + + )
{  fR = (X[j] - X[j - 1]) × (X[j] - X[j + 1]);
   if(fR > 0 && (X[j] - X[j - 1]) > 0) then the jth band is reflection apex;
   if(fR > 0 && (X[j] - X[j - 1]) < 0) then the jth band is absorption valley;
}
```

Here, N is band count, and $X[j]$ is albedo of the jth band.

After the location and related parameters are got, the detail analysis can be done. In general two methods are used, one is to give direct encoding and analysis to feature bands, and the other is to compute some quantitative index using feature bands and their parameters.

2.3 Encoding of spectral absorption or reflection features

The locations of feature bands are directly used in spectral feature encoding. The following will take absorption feature as an example. If one band is the location of absorption valley, its code will be "1", otherwise its code is "0". After the encoding is completed further matching and comparison can be done. Because of those uncertainties and errors in hyperspectral imaging process, the locations of feature bands perhaps move in near bands, and that will lead to low match ratio. In order to reduce the impact of band displacement, the extended encoding method is proposed and used in this paper. Its idea is that if the code of a certain band is "1" then the bands prior to and behind it will be assigned the same code "1", and then matching and analysis will be done.

The similarity measure to code vector is matching by bit. The matching ratio is got by the ratio of matched bands to total band count. In this study, two match schemes are used. One is matching the code of all bands and the other is only matching those feature bands.

Based on above analysis, four schemes are used and compared. These are: (1) direct encoding to all bands and matching by all bands, and (2) direct encoding to all bands and matching only by feature bands, and (3) extended encoding and matching by all bands, and (4) extended encoding and matching only by feature bands.

Table 3 is the results of analysis to three regions from hyperspectral image, and Table 4 is the results of analysis to data from USGS spectral database.

Table 3 Results of encoding to feature of data from hyperspectral image

	Scheme 1			Scheme 2			Scheme 3			Scheme 4			Combined encodingI		
	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
I	0.96	0.94	0.94	0.93	0.92	0.02	0.91	0.85	0.84	0.08	0.06	0.08	0.90	0.88	0.86
II		0.96	0.91		0.93	0.01		0.90	0.77		0.08	0.05		0.89	0.84
III			0.96			0.05			0.92			0.15			0.90

Table 4 Results of encoding to feature of data from USGS spectral database

	Scheme 1			Scheme 2			Scheme 3			Scheme 4			Combined encodingI		
	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
S1	1.00	0.82	0.67	0.11	0.02	0.02	1.00	0.71	0.40	0.30	0.15	0.21	1.00	0.69	0.40
S2		1.00	0.42		0.11	0.02		1.00	0.42		0.28	0.21		1.00	0.41
S3			0.67			0.28			1.00			0.73			1.00

From Table 3 and Table 4 we can know that when matching all bands are used, because there are many non-feature band, matching ratio of different regions is rather high. If matching by feature bands is used, the matching ratios of the same region is improved more effectively than that of different regions. But the matching ratio is low because feature bands are much less than total bands. After extended encoding is adopted, the match ratio of same region is higher than different regions and the feature band displacement problem can be solved to a great extent.

In order to analyze absorption and reflection features further, combined encoding is proposed. That means those absorption bands are encoded as "1", reflection bands as "2" and other bands as "0". Matching by bit is done after encoding, and the results are listed in the last three columns of Table 3 and 4.

From above analysis and comparison to spectral absorption and reflection feature encoding and matching, it can be found that although absorption and reflection band can describe the spectral properties of ground object, effective matching operation should be used in order to overcome the impacts of noise, band displacement and other factors. In practical applications, absorption and reflection can be used to extract thematic information and retrieve a certain type of object effectively.

Based on spectral absorption and reflection features, the spectral absorption index (SAI) or spectral reflection index (SRI) can be computed by wavelength, albedo of feature band and its left and right shoulders, and those indexes can describe spectral feature more precisely on some occasions.

3 Spectral computation and transformation features

Both correlativity and mutual compensation exist in different bands of hyperspectral RS information, so many new features can be got by certain computation and combination to some bands and used to classification, information extraction and other tasks.

3.1 Normalized difference of vegetation index (NDVI)

For hyperspectral RS information, NDVI can be viewed as a gradient function used to express the dramatic increase of vegetation ' s albedo at $\lambda_0 = 0.7\mu m$ location, and it can be expressed as Equation (1)

$$NDVI = \frac{X(\lambda_0 + \Delta\lambda) - X(\lambda_0 - \Delta\lambda)}{X(\lambda_0 + \Delta\lambda) + X(\lambda_0 - \Delta\lambda)} = \frac{1}{2X(\lambda)}\frac{dX}{d\lambda} \quad (1)$$

Here, $\lambda_0 = 0.7\ \mu m$, and $X(\ )$ is the albedo of corresponding band.

NDVI plays very important roles in hyperspectral application. It can describe some fine information about vegetation such as Leaf Area Index (LAI), ratio of vegetation and soil, component of vegetation and so on. In some classifiers (for example, ANN classifier) NDVI usually is used as an independent feature in classification.

3.2 Derivative spectrum

Derivative spectrum is also called as spectral derivative technique. One rank and two rank derivative spectrum can be computed by Equation (2) and (3)[4,5]

$$X'(\lambda_i) = \frac{X_{i+1} - X_{i-1}}{(\lambda_i - \lambda_{i-1}) + (\lambda_{i+1} - \lambda_i)} = \frac{X_{i+1} - X_{i\_}1}{2\Delta\lambda} \quad (2)$$

$$x''(\lambda_i) = \frac{X'(\lambda_{i+1} - X'(\lambda_{i-1})}{2\Delta\lambda} = \frac{X_{i+2} - 2X_i + X_{i-2}}{4\Delta\lambda^2} \quad (3)$$

Here, $\boldsymbol{X} = \{X_1, X_2, \cdots, X_N\}$ is the spectral vector of a pixel, and $\lambda = \{\lambda_1, \lambda_2, \cdots, \lambda_N\}$ is the wavelength set of a certain type of hyperspectral RS sensor.

Each rank derivative spectrum can be computed using algorithms similar to above. After derivative computation is end, we can find that each type of ground object may have some features distinguished from other entities in a certain rank derivative spectrum and that can be used to identify information. Sometimes derivative spectrum image can be used as the input of classifier directly. Although spectral derivative can provide new features in addition to original information, some new images will be formed after derivative operation and that will increase data volume dramatically. For M rank derivative spectrum, $N-2M$ bands will be formed, so how to process relationship between data volume and efficiency becomes a new question.

4 Spectral similarity measure features

In addition to analyze the shape of spectral curve and feature bands, direct similarity measure to spectral vectors is an effective method for hyperspectral information processing. Those similarity measure indexes in common use include spectral distance, correlation coefficient, spectral angle and spectral information divergence.

Spectral Angle (SA) has got wide application in hyperspectral information processing. SA is the angle of two pixel vectors with the same wavelength in spectral space, it can be computed by Equation (4)[3]

$$\cos\alpha=\frac{\boldsymbol{AB}}{|\boldsymbol{A}||\boldsymbol{B}|}=\frac{\sum_{i=1}^{N}\boldsymbol{A}_i\boldsymbol{B}_i}{\sqrt{\sum_{i=1}^{N}\boldsymbol{A}_i\boldsymbol{A}_i}\sqrt{\sum_{i=1}^{N}\boldsymbol{B}_i\boldsymbol{B}_i}} \tag{4}$$

Here, N is the count of band, $\boldsymbol{A}=(A_1,A_2,\cdots\cdots,A_N)$ and $\boldsymbol{B}=(B_1,B_2,\cdots\cdots,B_N)$ is two spectral vectors with the same wavelength set, $\boldsymbol{A}_i$ and $\boldsymbol{B}_i$ is albedo of the ith band of two pixels, α is spectral angle. In practice, it is unnecessary to compute α and its cosine can be used directly. The larger $\cos\alpha$ is, the more similar the two pixels are, otherwise the smaller $\cos\alpha$ is, the more different the two pixelsare are.

Spectral Information Divergence (SID) is also used to measure spectral similarity between pixels. It can be computed from equation (5)[6]

$$SID(A,B)=D(A\parallel B)+D(B\parallel A) \tag{5}$$

Here: $D(A\parallel B)=\sum_{i=1}^{N}p_i\log(p_i/q_i)$, $D(B\parallel A)=\sum_{i=1}^{N}q_i\log(q_i/p_i)$

$$p_i=A_i/\sum_{i=1}^{N}A_i,\ q_i=B_i/\sum_{i=1}^{N}B_i$$

In addition, the correlation coefficient of two pixel vectors can be computed by equation (6)

$$\rho=\frac{\sum_{i=1}^{N}(A_i-A_{\text{mean}})(B_i-B_{\text{mean}})}{\sqrt{\sum_{i=1}^{N}(A_i-A_{\text{mean}})^2}\sqrt{\sum_{i=1}^{N}(B_i-B_{\text{mean}})^2}}=\frac{Cov(\boldsymbol{A},\boldsymbol{B})}{Std(\boldsymbol{A})Std(\boldsymbol{B})} \tag{6}$$

The results of analysis to data from hyperspectral image and USGS spectral database are listed in Table 5 and 6.

Table 5 Results of spectral similarity measure indexes to data from hyperspectral image

	cosine of SA			SID			distance			correlation coefficientl		
	Ⅰ	Ⅱ	Ⅲ	Ⅰ	Ⅱ	Ⅲ	Ⅰ	Ⅱ	Ⅲ	Ⅰ	Ⅱ	Ⅲ
Ⅰ	0.990	0.960	0.539	0.114	0.360	14.659	0.306	0.977	0.510	0.979	0.905	0.746
Ⅱ		0.994	0.364		0.017	14.308		0.470	2.362		0.966	0.534
Ⅲ			0.930			42.682			0.274			0.977

Table 6 Results of similarity measure indexes to data from USGS database

	cosine of SA			SID			distance			correlation coefficientl		
	S1	S2	S3	S1	S2	S3	S1	S2	S3	S1	S2	S3
S1	1	0.99	0.90	0	0.01	0.24	0	0.32	13.0	1	0.99	0.57
S2		1	0.89		0	0.26		0	13.1		1	0.60
S3			1			0			0			1

From Table 5 and Table 6, it can be known that the cosine of spectral angle of a same pure region is much higher than that of different regions, so spectral angle is effective to measure the similarity between pixels and ground objects and used to classification, clustering and retrieval. At the same time, the SID of the similar regions is much less than that of different regions, so it is also useful. But distance and correlation coefficient are not very effective; especially distance can' t describe the similarity exactly. So spectral angle and SID are used more widely in hyperspectral RS information processing.

5 Conclusions and discussions

In this paper, oriented to the demands of hyperspectral RS information processing to spectral features, the framework of spectral features is proposed and some major feature extraction algorithms and their applications are discussed, and some improvement, experiments and analysis are finished. From the studies in this paper, the following conclusions can be drawn:

1) Based on the extraction principle and algorithm, spectral features in hyperspectral RS information can be categorized into three levels: spectral curve features, spectral transformation and computation features and spectral similarity measure features. This framework is useful for further analysis and applications.

2) As the common style of pixel spectral vector, some features can be extracted and used. The

algorithm and computation of binary encoding is simple and easy but it will lead to loss of some detail information. Quarternary encoding can describe curve features with high precision and be used to matching, retrieval and other work. The reflection and absorption features based on spectral curve have wide applications in retrieval, thematic information extraction and other tasks, but effective matching strategy must be adopted in order to control errors. In this paper two new approaches including extended encoding and matching and combined matching of reflectance and absorption features are proposed and it proved that they can get better results than traditional methods in feature measure.

3) As the main computation and transformation features, NDVI and derivative spectrum can provide new features participating in classification, extraction and other processing and extract those useful patterns and information hidden behind original data, so they are very useful in hyperspectral RS information processing.

4) For those spectra similarity measure indexes, Spectral Angle and SID are more effective than traditional indexes because they can measure the similarity more precisely, so they are usually used to classification, clustering and retrieval.

Some topics about the feature extraction and application of spectral feature are discussed in this paper. Our further studies will be focused on classification, object identification and thematic information extraction in hyperspectral RS information and the specific application modes of different spectral features in order to promote the development of hyperspectral RS application.

References

1 Pu Ruiliang, Gong Peng. Hyperspectral RS and its application. Beijing: High Education Press, 2000

2 Chen Shupeng, Tong Qingxi, Guo Huadong, *et al.* Study on RS information mechanics. Beijing: Science Press, 1999

3 Arel W, Michelle N, Brett B, *et al.* Spectral angle automatic cluster routine (SAALT): An unsupervised multispectral clustering algorithm. http://www.goole.com

4 Tsai F, Philpot W. A derivative-aided hyperspectral image analysis system for Land-Cover classification. *IEEE Transaction on Geoscience and Remote Sensing*, 2002, **10** (2): 416~425

5 Tsai T, William P. Derivative analysis of hyperspectral data. *Remote Sensing of Environment*, 1998, **66**: 41~51

6 Chang Chein-I. Spectral information divergence for hyperspectral image analysis. http://www.google.com

7 Du Peijun, Fang Tao, Tang Hong, *et al.* Similarity measure of spectral vectors based on set theory and its application in hyperspectral RS image retrieval. *Chinese Optical Letters*, 2003, **1**(11): 637~641

8 Wu Yan, Yang Wanhai, Li Ming. Fusion algorithm of multispectral and high-resolution panchromatic images. *Acta Photonica Sinica*, 2003, **32**(2): 174~178

9 Li Queyu, Wang Yang, Zhang Yonglin. In VIVD determinotion of the absorkance spectra of human skin. *Acta Photonica Sinica*, 2002, **31**(11): 1321~1325

10 Chen Zhenqiang, Shen Hongyuan, Zhang Ge, *et al.* Spectrum properties of hydrothermal emerald laser crystals. *Acta Photonica Sinica*, 2004, **33**(3): 382~384

高光谱遥感信息中的特征提取与应用研究

杜培军[1,2] 方 涛[2] 唐 宏[2] 陈雍业[2]

(1 中国矿业大学遥感与地理信息科学系,江苏 徐州 221008)

(2 上海交通大学图像处理与模式识别研究所,上海 200030)

收稿日期::2003-11-17

摘 要 特征提取、度量与分析是高光谱遥感应用的基础. 面向高光谱遥感数据的特点,将光谱特征划分为点尺度、面尺度和体尺度三个尺度的特征. 基于特征属性与算法原理,构建了光谱曲线特征、光谱变换特征和光谱度量特征三个层次的高光谱遥感光谱特征体系,并对光谱特征提取与应用进行了深入探讨. 光谱曲线特征包括直接光谱编码、光谱反射与吸收特征等,光谱变换特征包括植被指数、导数光谱等,光谱度量特征则包括光谱角、SID、相关系数和距离等. 在分析特征算法原理的同时对其特点和应用进行了探讨. 试验表明四值编码、光谱角和 SID 在应用中能够取得较好的效果.

关键词 高光谱遥感;光谱特征;特征提取;信息处理

Du Peijun (1975-), male, Ph. D., associate professor, postdoctor member of Shanghai Jiaotong University. His main research interests include remote sensing information intelligent processing and data mining.