Geoscience and Remote Sensing (2022)
Clausius Scientific Press, Canada

DOI: 10.23977/geors.2022.050102
ISSN 2523-2592 Vol. 5 Num. 1

Feature Recognition and Classification Based on Hyperspectral Data Mining

Mingliang Wang, Ying LI, Canghai Guan, Hongli Li

The Second Geodetic Surveying Brigade, Ministry of Natural Resources, Harbin, Heilongjiang 150025, China

Keywords: Hyperspectral, Data mining, Feature recognition and classification

Abstract: Since entering the 21st century, with the progress of various imaging spectrometer technologies and the rapid expansion of their application fields, hyperspectral remote sensing technology has been widely used in the fields of land and resources utilization, environmental monitoring, geological exploration, agricultural production, disaster early warning, urban planning, artificial target recognition and military detection. While providing more information than other remote sensing images, hyperspectral images have a large amount of data due to their many bands, which makes various ground object recognition and classification algorithms widely used in multispectral image processing slow and inefficient in hyperspectral image processing. Feature classification is an important field in digital image processing. A series of valuable data are obtained through the operation, extraction and classification of hyperspectral data images, and then applied in various fields. This paper studies the feature recognition and classification based on hyperspectral data mining. Hyperspectral data can accurately identify the common types of urban features, and the recognition method is particularly important for the results, and the fusion processing of hyperspectral images can improve the precision and accuracy of classification results to a certain extent.

1. Introduction

The development of high spectral resolution remote sensing technology is one of the major technological breakthroughs in earth observation made by human beings in the last two decades at the end of this century, and it is the frontier technology of remote sensing at present and even at the beginning of the next century. The images of the earth's surface obtained by hyperspectral imaging contain abundant triple information of space, radiation and spectrum [1]. Since the beginning of the 21st century, the progress of various imaging spectrometer technologies and the rapid expansion of their application fields have made hyperspectral remote sensing technology widely used in land and resources utilization, environmental monitoring, geological exploration, agricultural production, disaster warning, urban planning, artificial target recognition, military detection and other fields. In the field of remote sensing, hyperspectral remote sensing plays an important role, which makes the research of remote sensing technology to a new level [2]. In practical application, remote sensing can only be carried out in a limited band, such as ultraviolet, visible light, near infrared and so on. Medium-and wide-band materials are difficult to be detected in remote sensing, and hyperspectral

edge sensing and technology have solved this problem. In the late 1990s, with a series of basic problems of hyperspectral remote sensing application, such as calibration and quantification of hyperspectral imaging information, visualization and multi-dimensional expression of imaging spectral image information, image-spectral transformation, large data processing, etc., hyperspectral remote sensing has gradually shifted from the experimental research stage to the practical application stage [3]. Hyperspectral remote sensing images are rich in spectral information, image information and spatial information, so they have a very important application prospect in target recognition and feature classification.

With the increase of hyperspectral image band, the feature space dimension also increases. In this case, if the number of training samples is insufficient, the classification accuracy will rise first and then decline. This phenomenon is called hguhes phenomenon [4]. However, while providing more information than other remote sensing images, hyperspectral images have a large amount of data due to their many bands, which makes various ground object recognition and classification algorithms widely used in multispectral image processing slow and inefficient in hyperspectral image processing. Classifier type and classification scheme: People's subjective initiative will play a decisive role in this regard. For the same hyperspectral image, the selection of classifier type and classification scheme can make the classification results very different [5-6]. In this section, based on the brief introduction of previous achievements, I will put forward my feature optimized expert decision classification method. Combined with the recognition and classification ability of hyperspectral data and the spatial characteristics of high spatial resolution images, this paper makes full use of the advantages of two data sources and takes into account spectral and spatial information to carry out the research of feature recognition and classification.

2. Basic Concepts and Advantages of Hyperspectral Data Mining

Data mining refers to a multi-stage data processing process that uses modern information technology to extract effective information from massive data and form a knowledge pattern that can be understood by people. Specifically, in the field of remote sensing, hyperspectral data mining and knowledge discovery refer to extracting target feature information from hyperspectral images or high-resolution spectral curves through data processing and analysis, so as to acquire knowledge of mineral resources, ecological environment and social development [7]. Hyperspectral remote sensing, that is, high spectral resolution remote sensing, is a technology that can obtain hundreds of very narrow spectral band information and spectral continuous image data [8]. Hyperspectral image is a kind of high-dimensional image which contains spatial, episodic and spectral information and has huge data. On this basis, based on a large number of ground measured spectral information and pattern recognition technology, combined with the characteristics and related characteristic parameters of hyperspectral data, the development of spectral modeling, spectral matching, optimal band selection and geological mapping technology suitable for hyperspectral data processing and application is the focus and research hotspot of hyperspectral remote sensing in resource exploration at present. Hyperspectral data can obtain spectral characteristic curves of ground objects, and select or extract specific bands as needed to highlight target functions. The smaller the spectral angle between the image spectrum and its matched standard spectrum, the closer they are. Finally, according to the spectral angle matching results of different features, geological mapping is carried out by delineating the distribution intervals of features in the scatter map, as shown in Figure 1.

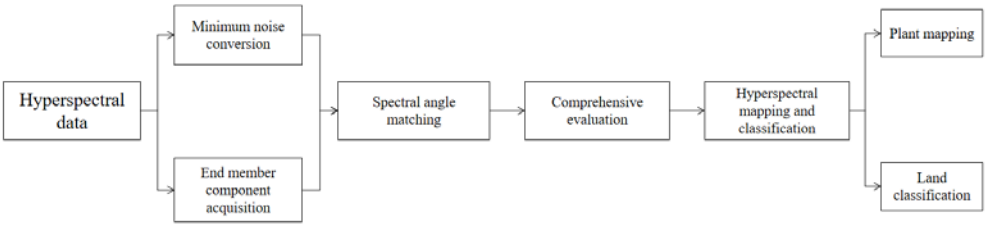

Fig.1 Work Flow of Hyperspectral Geological Mapping

In short, hyperspectral images combine the traditional spatial dimension and spectral dimension information, that is, “spectral image integration”, to form a coordinate system with wavelength as abscissa and gray value as ordinate, obtain the surface spatial information and generate a complete and continuous spectral curve, which makes it possible to invert the ground feature recognition according to the spectral characteristics. The main feature of hyperspectral remote sensing data is that it integrates the traditional image dimension and spectral dimension information, obtains the continuous spectral information of each feature while obtaining the surface spatial image, so as to realize the inversion of feature component information and feature recognition according to the spectral characteristics of features. Hyperspectral data is a cube of spectral image, which is composed of the following three parts

① Spatial image dimension: in the spatial image dimension, hyperspectral data is similar to general images. The general remote sensing image pattern recognition algorithm is a suitable information mining technology.

② Spectral dimension: a “continuous” spectral curve can be obtained from each pixel of hyperspectral image. The “spectral matching” technology based on spectral database can realize the purpose of identifying ground objects.

③ Feature space dimension: hyperspectral images provide a super dimensional feature space. For Hyperspectral Information Mining, it is necessary to deeply understand the distribution characteristics and behavior of ground objects in the n-dimensional feature space formed by hyperspectral data. It is found that the hyperspectral high-dimensional space is quite empty, the data distribution is uneven, and tends to be concentrated at the corner of the super dimensional cube space. Due to the differences of typical data, It can be mapped to a series of low dimensional subspaces, so there is an urgent need to develop effective feature extraction algorithms to find low dimensional subspaces that maintain important differences, so as to effectively realize information mining.

3. Feature Recognition and Classification

3.1 Feature Recognition and Classification Based on Hyperspectral Data Mining

After decades of development, the processing and analysis technology of hyperspectral remote sensing has made great progress. Based on the traditional classification algorithm, a series of classification methods for the characteristics of hyperspectral data have been summarized and formed. Compared with the traditional supervised classification and unsupervised classification, Using the known spectral data in the spectral database and using the matching algorithm to identify the ground coverage type in the image is the most characteristic. Feature classification is an important field in digital image processing. A series of valuable data are obtained through the operation, extraction and classification of hyperspectral data images, and then applied in various fields. Remote sensing image feature classification has also become the most important task in the application of hyperspectral remote sensing. The biggest difference between hyperspectral image

target recognition and hyperspectral image classification is that whether supervised classification or unsupervised classification, the sample spectrum of the category comes from the image itself. For unsupervised classification, it is necessary to assume the possible number of categories in the image region in advance, and then complete an ideal region division in the pixel spectral space based on the mathematical model. A great difference between hyperspectral image data and multispectral image data is that the spectral bands of the former contain specific physical meaning. The understanding of ground objects cannot stay on the basis of mathematical statistics of samples. We should start with ground spectral analysis of ground objects to optimize spectral characteristics and construct exclusive spectral characteristic parameters. The construction of spectral characteristic parameters should have the first priority in the selection of classification algorithm. However, through reading the literature, it is found that through practical attempts, the minimum distance matching method between spectra is sensitive to noise. The ground object category of hyperspectral remote sensing image at urban scale is more complex than that of mining area, and the noise of image is more obvious. Therefore, the minimum distance method is not ideal to obtain the ground object classification, and this method cannot debug the results in practical operation, The error is large.

3.2 Mixed Spectral Classification Method

In the remote sensing image, the image pixel records the sum of the target's blessing energy in the ground range corresponding to the instantaneous field of view angle of the detection unit. If this ground range only contains targets of the same nature, the pixel records the ground targets of the same nature, and such pixel is called “pure pixel”. If the ground range corresponding to the instantaneous field of view angle of the detection unit contains many kinds of targets with different properties, and it records the synthesis of spectral signals of various types of ground objects, this pixel is called “mixed pixel”. In the remote sensing program, the enhanced spectral resolution produced by current sensing technology brings new potential and challenges to data analysis. On the one hand, the availability of a large number of spectral bands makes it possible to identify more specific categories more accurately, which is impossible for earlier sensor data to complete. On the other hand, a large number of available interrelated classes and spectral bands need a lot of training samples. Unfortunately, these training samples are expensive and difficult to obtain. Therefore, classification statistics must be estimated from a limited set of training samples. The standard nearest neighbor classification is the supervised classification based on Yikang software platform. Firstly, the algorithm defines the categories of image objects, selects representative samples, and calculates the features of the sample categories to complete the classification of images in a feature space. The fuzzy model is based on the fuzzy set theory. Unlike the concept of classification, a pixel is not definitely classified into a certain category, but is associated with more than one class at the same time. Its basic principle is to treat all kinds of ground objects as fuzzy sets. Pixels are elements of fuzzy sets. Each pixel corresponds to a set of membership values, and the membership degree represents the area percentage of this ground object category contained in pixels.

4. Conclusions

Hyperspectral data mining and knowledge discovery technology is one of the frontiers of modern remote sensing, and it is also a research hotspot. With the continuous intensification of urbanization, people pay more and more attention to the monitoring of urban environmental changes. Because of the influence of human factors on the types of urban objects, the spectral characteristics are complex and diverse, and the traditional multispectral remote sensing data lacks spectral information, it is difficult to finely classify and identify urban objects. Hyperspectral images can

obtain detailed spectral information of ground objects and become a powerful tool for urban remote sensing research. Based on the diagnosable spectral information of ground objects, the spectral information of strata and background ground objects in the study area is directly extracted from hyperspectral images through high-precision data preprocessing, hyperspectral data mining, spectral knowledge discovery and geological mapping technology research, thus providing good technical support for stratum identification research using hyperspectral remote sensing technology. If multiple classifiers can be combined by multi-kernel method, each classifier can fully exert its respective advantages and avoid their disadvantages, and further improve the accuracy of hyperspectral classification. Multi-kernel learning has important practical significance and application prospect in hyperspectral image classification.

References

[1] Hai-Hong P, Ngoc-Son V, Vu-Lam N, et al. Action recognition based on Motion of Oriented Magnitude Patterns and Feature Selection[J]. IET Computer Vision, 2018, 12(5):735-743.

[2] Cai J, Hu J, Li S, et al. Combination of temporal-channels correlation information and bilinear feature for action recognition[J]. IET Computer Vision, 2020, 14(8):634-641.

[3] Huthaifa, I, Ashqar, et al. Smartphone Transportation Mode Recognition Using a Hierarchical Machine Learning Classifier and Pooled Features From Time and Frequency Domains[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 20(1):244-252.

[4] Sl A, Bz A, Sz A, et al. Local discriminant coding based convolutional feature representation for multimodal finger recognition - ScienceDirect[J]. Information Sciences, 2021, 547:1170-1181.

[5] Bi L, Xie W, Zhao J. Automatic recognition and classification of multi-channel microseismic waveform based on DCNN and SVM[J]. Computers & geosciences, 2019, 123(FEB.):111-120.

[6] Appice A, Guccione P, Malerba D. A novel spectral-spatial co-training algorithm for the transductive classification of hyperspectral imagery data[J]. Pattern Recognition, 2017, 63:229-245.

[7] Zhao S, Zhang B, Chen C P. Joint deep convolutional feature representation for hyperspectral palmprint recognition[J]. Information Sciences, 2019, 489:167-181.

[8] Pg A, Gqb C, JK D. Discriminant analysis and feature selection in mass spectrometry imaging using constrained repeated random sampling - Cross validation (CORRS-CV)[J]. Analytica Chimica Acta, 2020, 1097:30-36.