

IRON ORE DETECTION BY REMOTE SENSING TECHNOLOGY BASED ON PRINCIPLE COMPONENT ANALYSIS AND SPECTRAL ANGLE MATCHING

Pengfei Pan¹, Houguang Sun¹, Chunyang Li¹, Yachun Mao², Shanjun Liu², Dong wang²

¹Anqian Mining Limited Liability Company of Anshan Iron and Steel Group Corporation, Anshan, China

² College of Resources and Civil Engineering, Northeastern University, Shenyang, China

Email: liusjdr@126.com

KEY WORDS: Hyperspectral remote sensing, Iron ore, Principle component analysis, Spectral angle, Mine detection

ABSTRACT: Iron ore is the main mineral resources to support China's national economy. The magnetite and hematite are the main types of iron ore, which accounts for 73.5% of the total amount. Traditionally, the way of iron ore detection is from geological survey, geophysical prospecting and drilling, which is impossible to acquire information at large scale. Remote sensing technology provides the way with speedy and efficacy in mine detection at large scale. Additionally, the hyper-spectral remote sensing has the advantage in geological exploration and mine detection with the comprehensive spectral information. This study aims to establish possible remote sensing methods for iron ore detection by field spectroradiometer technology and mathematical models. The samples including magnetite and hematite were selected from Anqian iron mine in the northeastern China. The samples are processed to thickness of 1-2 cm with the circular shape by drilling and cutting, and the chemical compositions and spectrum were measured. The principle component analysis (PCA) method is conducted to reduce dimension of huge spectral data. Furthermore, spectral angle method is used to identify and classify iron ore types by establishing spectral library of magnetite and hematite. The results showed that the PCA can capture the maximum and major information within the data by reducing dimension of comprehensive data; it keeps the significant information for spectrum regions by ensuring the indicator is independent, which is useful for further remote sensing model establishing. The spectral angular method performed well in the classification iron ore with accuracy up to 95%. This method has potentiality and possibility to be applied in mine detection using hyperspectral remote sensing technology.

1. INTRODUCTION

Iron ore is the main mineral resource to support China's national economy, while the magnetite and hematite are the main types of iron ore, accounting for more than 70% of the total ore in China. In the past, iron ore exploration depended mainly on field geological survey, geophysical prospecting, drilling and other methods, belonging to the point measurement and not able to get a large area of information. The rise of hyperspectral remote sensing technology provides a convenient and efficient way for iron ore exploration of large area, and it has more advantages in geological exploration, because of the high spectral resolution[1-3]. Arne Grumpe et al have utilized hyperspectral data to achieve the moon mineral mapping from shading and reflectance [4]. Yongbin Sun et al. have solved the alteration parameters to establish the alteration Remote Sensing Model for prospecting using hyperspectral data [5]. Although hyperspectral remote sensing has made a lot of progress in the field of geological prospecting, problems do exist with this approach-large amount of data and complex data processing to cause a lot of inconvenience.

The aim of this study is to develop a method for the remote sensing detection of iron ore via measuring the spectra of samples and mathematical analysis. The first thing is to gather magnetite and hematite samples from Anqian mine belonging to Anqian Mining Limited Liability Company. And then a circular sheet of 1-2 cm thick is formed through hole drilling and cutting, after which chemical compositions and visible/near infrared spectra of the samples are determined and measured. In order to simplify the processing of hyperspectral data process, the principal component analysis is applied for dimensionality reduction of hyperspectral data, and then uses the spectral angle matching method to match the target spectrum for realizing remote sensing identification of iron ore. Finally, the experimental verification is made by using the field samples.

2. PROPOSED METHODS

Numerous classification methods can be applied to hyperspectral remote sensing, of which spectral matching technique has become the relatively prominent approaches, referring to a method for matching unknown ground-object spectra using reference spectra in the spectral library, identifying the type of the corresponding surface features, and then to achieve the purpose of object recognition, on the basis of the similarity between the reference and the unknown spectrum.

Due to the high spectral resolution, there are many restrictions to utilize the hyperspectral data by spectral matching

method for remote sensing identification, such as the great amount of data, not conducive to the calculation and data redundancy [6]. This paper introduces the optimization algorithm that dimension of the hyperspectral data should be reduced firstly, and then the spectral angle matching is carried out to determine the categories of samples, based on the descending dimension spectral data.

2.1 Principal component analysis

The principal component analysis was first put forward by Hotelling in 1933 and it is a multivariate statistical analysis method which transform the multiple indicators into a few comprehensive indexes by use of dimension reduction [7]. The principle is to calculate the characteristic values of the multi-index vector and select several larger values according to a certain rule, and then project the sample points onto the selected feature vector to realize the dimension reduction of the multi-index vector.

Assuming that a n -dimensional vector include multiple indexes- x_1, x_2, \dots, x_n , and they are turned into a m -dimensional principal component vector- y_1, y_2, \dots, y_m , and moreover, m is less than n , they will constitute the following formula.

$$\begin{cases} y_1 = l_{11}x_1 + l_{12}x_2 + \dots + l_{1n}x_n \\ y_2 = l_{21}x_1 + l_{22}x_2 + \dots + l_{2n}x_n \\ \vdots \\ y_m = l_{m1}x_1 + l_{m2}x_2 + \dots + l_{mn}x_n \end{cases} \quad (1)$$

Coefficient l_{ij} ($i=1,2, \dots, m$; $j=1,2, \dots, n$) in the formula above can be obtained from the score matrix through the analysis of the principal components.

A principal component analysis model can be described by three attributes to describe the principal component score, the principal component load, and the explained variance. Several samples of similar scores on a principal component means that the samples are similar in composition and vice versa. The principal component load is used to describe the correlation coefficient between the principal component and the original variable. The larger the correlation coefficient is, the more appropriate the principal component is to the representativeness of the variable. Explained variance is used to indicate the interpretation percentage of the principal component to the original data. By the analysis of principal component, the influence of spectral transformation process on the structure of spectrum data can be intuitively understood, which is convenient to simplify the matching process of spectral analysis and matching in the mathematical perspective.

The method for processing hyperspectral data can effectively identify the main band information, remove redundant information, and reveal the simple structure hidden behind the complex data by reducing the dimension of the original dense band data, which not only retains the main information of various bands but also ensures that the comprehensive indexes are not related to each other and contributes to the establishment of remote sensing detection model.

2.2 Spectral angle matching

Spectral angle matching is a common method in hyperspectral remote sensing ground-objects identification. This method is also called spectral angle mapping method, which compares the target spectrum with standard spectra (reference spectra) to obtain the generalized included angle between two spectral vectors. And the smaller the angle, the greater the similarity between the two objects. When the angle is smaller than a certain threshold value, it is considered that the target spectrum is properly matched with the standard spectrum.

If the standard principal component score vector is T and the reference principal component score vector is R , while T and R are not zero vectors, their generalized included angle θ will be defined as the formula (2):

$$\theta = \arccos \frac{T \cdot R}{|T| \cdot |R|} \quad (2)$$

Spectral angle matching rules is described that the narrow spectral angle threshold value set, when the narrow spectral angle is smaller than the value, the target spectrum is considered to matching with the reference spectrum successfully, otherwise the match is not successful.

3. METHOD VALIDATION

3.1 Sample collection and standard spectra establishment

The experimental samples are collected from Anqian Mining Limited Liability Company in Liaoning province of China, which is a large open-pit mine and whose mining orebody is the typical metamorphic iron ore. The surrounding rocks mainly consist of old metamorphic and migmatitic granite from the Archean to early Proterozoic, including chlorite schist, mica schist, phyllite and amphibolite, etc. There are two main types of iron ore- hematite (chemical composition of Fe_2O_3) and magnetite (chemical composition of Fe_3O_4), respectively.

The study captures the spectra of samples by SVC HR-1024 portable spectrometer, whose wave band range is $0.35\sim 2.5\mu\text{m}$, channel number is 1024, spectral resolution is less than or equal to 8.5 nm and view angle is 4° . The spectral tests are conducted in 10:00-14:00 to reduce the impact of aerosols, when the sky is cloudless and sunny, and the sun elevation angle is at about 45° . When measuring, the surface of the samples are kept level and the spectrometer lens is perpendicular to the observation surface, consistent with the observation of the satellite sensor. The sampling integral time is set to 3 seconds, while each sample receives repeated observation twice, and the average values are adopted. The results of spectral test for all samples are shown in figure 1.

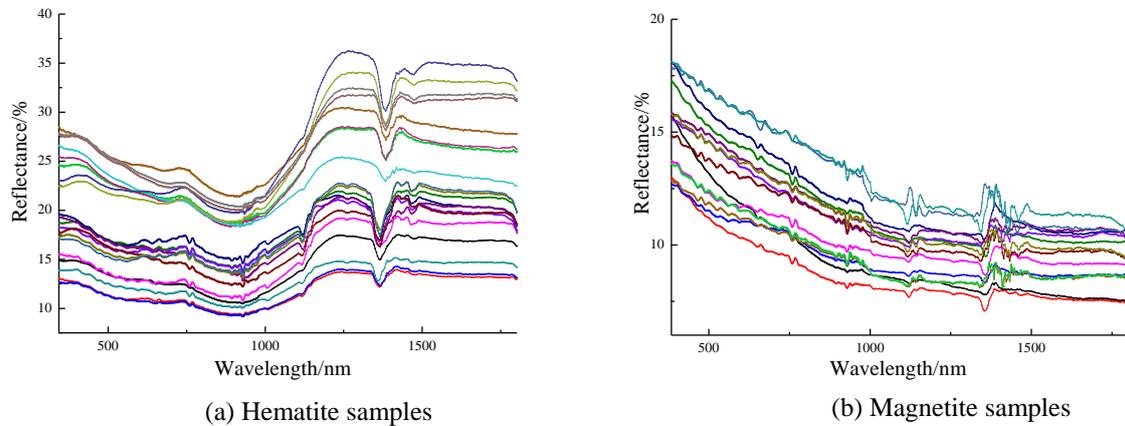


Figure 1 Spectral curves of the experimental samples

Figure 1 (a) shows that the spectral characteristics of hematite samples takes on a relatively gentle downward trend in $343\text{nm}\sim 978\text{nm}$ band, an obvious upward trend in the $978\text{nm}\sim 1255\text{nm}$ band, and the overall moderate trend after 1255nm band, in which it appears to be two significant wave valleys in 1121 nm and 1380 nm band, and the depth of the trough at 1121 nm is below than that at 1380 nm .

Illustrated from Figure 1 (b), spectral characteristics of magnetite samples present an overall decline trend in the $343\text{nm}\sim 1801\text{nm}$ wavelength range, and shows a gentle trend after 1473nm . Similarly, there are also valleys around the 1121nm and 1380nm band, but the depth of trough is less than those of hematite.

The main difference between the spectral characteristics of hematite and magnetite can be found:

In the wavelength range of $343\text{nm}\sim 978\text{nm}$, the downward trend of spectral curves for magnetite samples are more obvious than hematite. However, in the $978\text{nm}\sim 1255\text{nm}$ band, spectral curves of hematite appear to rise, and the spectral curves of magnetite remain downward trend. In addition, the spectral curves of magnetite and hematite samples have appeared a trough at 1121nm and 1380nm , but the trough depth of hematite spectral curves at 1380nm is greater than that of magnetite.

3.2 Principal component analysis processing

In this study, the spectra data for the 2 kinds of Anqian mine samples have been processed with the principal component analysis, and the extracted principal components can be calculated by regression prediction equation^[8]. Its procedures can be divided into the following three steps:

- (1)The correlation matrix is calculated from the spectral data of samples as well as judging their regression and co-linearity
- (2)The characteristic value and vector are calculated by the correlation matrix and sorted to select the main factors and determine the main component

(3)The principal component contribution rates are calculated and the extraction effect is analyzed. Finally, several principal components of spectral data can be calculated,according to the results of the steps(2).

After principal component analysis processing mentioned above, the 3 principal components are finalized based on the principle of over one eigenvalue, which represent the entire spectrum information and its principal component eigenvalue and contribution rate of variance are shown in table 1. It is seen from table 1 that, through principal component analysis, original 703 bands information can be replaced by 3 principal components, and the cumulative contribution rate is up to 99.766%. That is to say, 3 principal components covers over 99% information of the original band information, which can represent the features of spectral curves.

Table 1 Characteristic value and variance contribution rate of principal component analysis

Component	Characteristic value of correlation matrix		
	Characteristic value	Variance contribution rate (%)	Cumulative contribution rate (%)
1	645.140	91.770	91.770
2	47.795	6.799	98.568
3	8.418	1.197	99.766

3.3 Spectral angle matching processing

Firstly, reference spectra must be determined to realize spectral angle matching. Thus the average value of 3 principal components should be calculated and synthesized the sample data as a reference spectral sample data, in order to get the reference spectra of hematite and magnetite. And then, the spectral angles should be calculated by all the 20 verification samples' spectra and the 2 reference spectra to judge the types of samples.

2 reference spectral vector of principal components can be expressed by the following equation: $E_i=(e_{m1}, e_{m2}, \dots, e_{mj}), (m=1,2 ; j=1,2,3)$. When $m=1$, the result means the reference vector of principal components for hematite. When $m=2$, it means the reference vector of principal components for magnetite. In the formula above, j represents the serial numbers of principal components, such as the result means the first principal component when $j=1$.

If check vector of 3 principal component for each verification sample is expressed as $C=(c_{p1}, c_{p2}, \dots, c_{pj}), (p=1,2, \dots, 20 ; j=1,2,3)$, in which p refers to the serial numbers of validation samples, and the result represents those of hematite samples when p is in the range from 1 to 10, the same as for magnetite samples when p is in the range from 11 to 20, the calculation formula of the narrow spectral angle will be as follows:

$$\theta_p = \arccos \frac{\sum_{j=1}^3 e_{mj} c_{pj}}{\sqrt{\sum_{j=1}^3 e_{mj}^2} \sqrt{\sum_{j=1}^3 c_{pj}^2}}, \theta_p \in \left[0, \frac{\pi}{2} \right] \quad (3)$$

The narrow spectral angles of the verification samples and reference samples, hematite and magnetite, can be respectively represented by θ_{He} and θ_{Ma} . After all spectral data of samples processed, the results of narrow spectra angles are shown in Table 2 and Table 3.

To realize the matching between the spectra of verification and reference samples, the study formulates the following rules: First of all, the narrow spectral angle threshold should be set to of 10 degrees, namely when the narrow spectral angle is less than 10° , which means that the verification spectrum and the reference spectrum are successfully matched, otherwise, not successfully. Second, the verification sample containing smaller spectral angle should be selected as the sample matched successfully, when there are 2 verification samples can be matched with the same reference sample.

From table 2 and table 3, all the 10 hematite samples are matched successfully after spectral angle matching processing and the accuracy is 100%, while 9 of 10 magnetite samples have been matched correctly, and the correct rate is 90%. The matching correct ratio of the total samples is 95%, indicating the method has a good feasibility to determine the ore type of the samples.

Table 2 Calculated narrow spectral angles (°) of 10 hematite samples

Verification sample	θ_{He}	θ_{Ma}	Judgment result
Hematite -1	2.21	8.86	Hematite
Hematite -2	8.13	19.00	Hematite
Hematite -3	7.79	18.67	Hematite
Hematite -4	3.37	14.06	Hematite
Hematite -5	3.36	13.99	Hematite
Hematite -6	1.22	10.78	Hematite
Hematite -7	0.93	10.67	Hematite
Hematite -8	4.64	6.41	Hematite
Hematite -9	2.76	8.16	Hematite
Hematite -10	6.86	17.70	Hematite

Table 3 Calculated narrow spectral angles (°) of 10 magnetite samples

Verification sample	θ_{He}	θ_{Ma}	Judgment result
Magnetite -1	9.85	1.62	Magnetite
Magnetite -2	5.28	5.87	Hematite
Magnetite -3	10.01	0.87	Magnetite
Magnetite -4	9.99	0.95	Magnetite
Magnetite -5	10.75	1.23	Magnetite
Magnetite -6	11.34	0.92	Magnetite
Magnetite -7	12.14	1.96	Magnetite
Magnetite -8	11.61	1.59	Magnetite
Magnetite -9	10.28	1.15	Magnetite
Magnetite -10	10.98	0.17	Magnetite

4. CONCLUSIONS

In view of the large amount of data and complicated data processing for hyperspectral remote sensing detection in iron ore mine, the study has introduced the principal component analysis method to take effective data dimensionality reduction, and then exploited the method of spectral angle matching for the Magnetite and Hematite identification and classification. Finally, verification experiments have been conducted by measuring the samples' spectra of Anqian Mine. The main conclusions of this study are as follows:

(1)The proposed method can effectively simplify the data processing program. There are 703 bands in the original spectral data obtained by spectrometer measurement, which are simplified to 3 indexes after principal component analysis. Base on the spectral angle matching with these 3 indexes, we can greatly simplify the calculation process, reduce the amount of computation, optimize the hyperspectral data processing, and minimize the loss of spectral information.

(2)The proposed method has been verified by the samples from Anqian Mine, and the total success rate is 95%, which shows this method has better applicability and operability in the iron ore detection by hyperspectral remote sensing.

REFERENCE

- [1] Nandibewoor A, Hegadi R, Adiver P., 2014. Identification of vegetation from satellite derived hyper spectral indices[C], 2014 International Conference on Contemporary Computing and Informatics, pp.1215-1219.
- [2]Picon A, Ghita O, Rodriguez-vaamonde S, *et al.*, 2011. Biologically-inspired data decorrelation for hyper-spectral imaging [J]. EURASIP Journal on Advances in Signal Processing, pp.66:1-10.
- [3]Liu Jianjun, Wu Zebin, Sun Le, *et al.*, 2011. The Design and Implementation of Hyperspectral Remote Sensing Data-Processing System Facing Mineral Identification[C], International Conference on Remote Sensing, pp.156-159.
- [4]Grumpe A, Zirin V, Wohler C., 2015. A normalisation framework for (hyper-) spectral imagery [J]. Planetary and Space Science, pp.111:1-33.
- [5]Sun Yongbin, Li Mingsong, Wang Ruijun, *et al.*, 2014. Deposit location determination based on hyperspectral alteration information from airborne hyperspectral images and metallogenic geologic environment [J]. Gold, 12 (35), pp.8-13.
- [6]Zhang Da, Zheng Yu-quan., 2013. Hyperspectral Remote Sensing and Its Development and Application Review[J].

Optics & Optoelectronic Technology, 11 (3) , pp.67-73.

[7]Hotelling H., 1933. Analysis of a complex of statistical variables into principal components [J]. Journal of Educational Psychology, 24 (6) , pp.417-441.

[8]Lu Yan-li, Bai You-lu, Yang Li-ping, *et al.*, 22008. Application of hyperspectral data for soil organic matter estimation based on principle components regression analysis [J]. Plant Nutrition and Fertilizer Science, 14 (6) , pp.1078-1079.