Research on Hyperspectral Surface Reflectance Dataset of Typical Ore Concentration Area in Hami Remote Sensing Test Field

Shuneng Liang 1, Yang Li 2, Hongyan Wei 1, Lina Dong 1, Jiaheng Zhang 3, Chenchao Xiao 1

1 Land Satellite Remote Sensing Application Center, Ministry of Natural Resources, Beijing 100048, China;
2 Key Laboratory of Education Ministry of Precision Opto-mechatronics Technology, School of Instrumentation Science and Opto-electronics engineering, Beihang University, Beijing 100191, China;
3 Institute of Earth Sciences and Resources, China University of Geosciences (Beijing), Beijing 100083, China

Key words: Hyperspectral Remote Sensing, Hami Remote Sensing Test Field, Surface Reflectance, Dataset, Ore Concentration Area.

Abstract

Surface reflectance data is the basic data source for the hyperspectral parametric remote sensing products and remote sensing quantitative application, which is widely used in various application fields such as natural resources and ecological environment monitoring. At present, multispectral data takes the leading role among the common land surface reflectance datasets and the reflectance data mainly involves the types of ground objects such as farmland, forest land, water body, soil, etc., while the datasets relatively less targets the types of rock and mineral surface objects, yet especially the reflectance datasets with the combination of time series and multi-scale satellite-earth are even more scarce. In order to better promote the application of hyperspectral surface reflectance and explore the advantages of joint application of satellite-earth multi-scale reflectance data, on the basis of field-measured rock and mineral target spectral, a comprehensive surface reflectance dataset was generated by using domestically produced hyperspectral satellite data as the data source in this study, mainly focusing on the typical ore concentration area in the Hami Remote Sensing test field in Xinjiang. The dataset includes multi-period hyperspectral satellite surface reflectance images, field measured rock and mineral spectral data, and multi-period sub-pixel spectral data collected based on ground spectral measured points, which can provide significant support for the research and development and accuracy verification as well as performance evaluation of algorithms such as surface reflectance inversion, mineral identification and ground object classification.

1. Introduction

Surface reflectance is the essential attribute of surface objects, and it is the most basic remote sensing surface parameter and one of the effective factors to distinguish the categories of surface objects (Wen Jianguang et al., 2015; Lin Xingwen et al., 2020; Cao Xiaojie et al., 2019). As the most important quantitative remote sensing products, surface reflectance data is the basic data source of many parametric remote sensing products, which is of great significance for quantitative remote sensing applications and is widely used in typical application fields such as forestry, agriculture, water resources, ecological environment and urban environment (Liu Liangyun et al., 2017; Cao Xiaojie, 2018; Shen Qian et al., 2021).

Up to now, the remote sensing datasets which have been made public and widely used mainly include two types: multispectral and hyperspectral dataset. Multispectral datasets can be divided into long-series medium-resolution multispectral datasets and high-spatial resolution multispectral datasets. Common long-series medium-resolution multispectral datasets include reflectance products such as MODIS, Landsats and Sentinel-2 (Vermote et al., 2016; Richter and Schlapfer, 2015; Zhang Zhaoaming et al., 2020; Peng Yan et al., 2020; Cao Xiaojie, 2019; Doxani et al., 2018). Common multi-resolution datasets with high spatial resolution include UC-Merced dataset (Yang and Newsam, 2010), WHU-RS dataset (Xia et al., 2010), AID dataset (Xia et al., 2010), the NPWU-RESISC dataset, the EuroSAT dataset (Helber et al., 2019), the RSSC7N (Zou et al., 2015), and the reflectance datasets of GF-2 in the plain area of Beijing (Shen Qian et al., 2021). Common hyperspectral datasets, including Indian Pines dataset, Pavia University dataset, Salinas dataset, Pavia Center dataset, KSC dataset, Cuprite dataset, Botswana dataset (Xu Min et al., 2017; Li Zhuqiang et al., 2018; Huang et al., 2017; Zhang et al., 2012; Li et al., 2016) and the domestic tea tree dataset of Changzhou City, Jiangsu Province, TGHRSSC dataset and the dataset of Majiwan Village, Xiongan New Area (Zhang et al., 2017; Liu Kang et al., 2020; Cen Yi et al., 2020), are mostly acquired by airborne photography, and the data phase and scale of data acquisition are relatively single. Common multispectral datasets with high spatial resolution are mainly used for fine classification and ground object. Compared with the wide application of hyperspectral remote sensing datasets in European and American countries, the publication and application of hyperspectral remote sensing datasets in China are still less. In recent years, with the launch of China's GF-5, 5m Optical satellites and other series of hyperspectral satellites as well as the construction and application of remote sensing geological test fields, China's ability to acquire high-quality hyperspectral remote sensing data has been greatly enhanced, which has greatly improved the types, quantity and quality of China's hyperspectral remote sensing data sources and provides significant support for promoting China's hyperspectral remote sensing application research and operational capabilities (Tong Qingxi et al., 2016; Wang Yueming et al., 2016; Zhang Xia et al., 2017; Zhao et al., 2016; Liang Shuneng et al., 2015; Li Shengyang et al., 2019).

In this study, taking the typical ore concentration area of Xinjiang Hami Remote Sensing Geological test field (hereinafter referred to as Hami test field) as the study area, a comprehensive dataset of hyperspectral surface reflectance in typical ore concentration area of the test field was constructed, based on multi-period domestic hyperspectral satellite images, and using the pre-processing technologies of radiation correction and atmospheric correction, and combining the spectral data of ground rock and ore with atmospheric parameter data obtained from field measurement in the early construction of the test field. Compared with the existing datasets, this dataset, characterized by high spectral resolution, multiple spectral categories and scales of rock and ore surface objects, and the complete matching parameters of spectral data, can provide good data support for the research and development of hyperspectral application technology, and serve...
the design and demonstration of hyperspectral remote sensing payload and the authenticity inspection and verification of quantitative remote sensing products.

2. General Introduction of the Study Area

2.1 Overview of Physical Geography

Located in the southeast of Hami City in the East Tianshan Mountains of Xinjiang, about 160km away from the urban area of Hami.

![Figure 1. Geographical location of hami test field](image)

The Hami test field is situated in Gobi Desert terrain, with characteristics of continental climate, lack of rainfall, underdeveloped water system and no perennial running water. There is a large temperature difference between day and night, hot in summer and cold in winter, sparse vegetation and no permanent residents. Ore industry is relatively well-developed with large or larger mining bases having been built in the Huangshandong copper-nickel deposit area. Regional geographical conditions are suitable for the research and application of remote sensing technology.

![Figure 2. Natural physiographic characteristics of test field](image)

2.2 Regional Geological Characteristics

The test field area is located in the Junggar block of the Kazakhstan-Junggar plate, and the regional structure is well-developed, which is relevant to the multiple stages and complexity of the tectonic movement in the area. This is related to the multi-stage complexity of tectonic movement in the region. Deep faults, major faults and general faults constitute a complex fault system, which divides the geotectonic system with different geological construction characteristics. Intrusive rocks and dike rocks are well developed in the area, ranging from ultrabasic and basic rocks to medium acid rocks. Affected by supergenic geological process, intrusive rocks in the area are strongly altered, among which the limonization, chloritization, aging and pyritization of medium acid rocks are well developed (Liang Shuleng et al., 2015).

3. The Research and Development Methods of Dataset

3.1 Data Source

The hyperspectral satellite data used in this paper are mainly 20 scenes of L1A-grade products of Gaofen-5 (hereinafter referred to as GF-5) and Ziyuan 1-02D (hereinafter referred to as ZY1-02D) covering research areas from January 2019 to June 2022. The characteristics of data parameters are shown in Table 1. (Liu Yinnian et al., 2018; Sun Yunju et al., 2019; Liu Yinnian et al., 2020; Liang Deyin et al., 2020).

![Table 1. Characteristics of hyperspectral satellite data](image)

3.2 Data Acquisition and Processing Methods

3.2.1 Production of Hyperspectral Satellite Surface Reflectance Product: 20 scenes of good quality hyperspectral images (L1A-grade) covering the study area with cloud cover less than 15% were selected from GF-5 and ZY1-02D hyperspectral satellites in transit through the study area from January 2019 to June 2022. The original L1A-grade data of hyperspectral satellite are DN value data with PRC correction parameters, which are processed by band synthesis, radiation correction and atmospheric correction (Dong Xinfeng et al., 2020; Liu Yao et al., 2022).

![Figure 3. Comparison of hyperspectral data cube and typical surface reflectance in study area](image)

First, the data of VNIR and SWIR bands are synthesized, and overlapping bands of shortwave infrared are eliminated. Secondly, the DN value data is converted to radiance data through radiometric calibration. Finally, the surface reflectance image of hyperspectral satellite in the study area is generated by FLAASH atmospheric correction program.
Based on the work above, some control data and DEM are used to conduct orthorectification to generate a single field ortho-corrected reflectance data product. The water vapor absorption bands near 1400nm and 1900nm can be further removed as needed during the data use.

### 3.2.2 Field Measurement of Ground Object Spectra

Based on the main rock types, altered minerals, and ore-bearing rock masses, surrounding rocks, and mineralization alteration phenomena of strata and rock masses exposed in the study area, spectral measurement profiles were arranged according to the requirements of rock and mineral spectral measurement technology to obtain spectral data of various main types of rocks and ore in the study area (Liang Shuneng et al., 2015). Based on the well-arranged spectral measurement profile, a total of 468 measurement points were set up along the profile line according to the changes in lithology and altered minerals of the rock mass. In accordance with the standard for rock and mineral spectral testing (Wang Runsheng et al., 2011; Yan Baikun et al., 2014) and the technical requirements for the construction of the test field, the portable spectral Radiometer (ASD field spectrometer) was used to measure not only the spectral data at each measuring point, but also the rock and mineral spectra of four auxiliary points within 15-20m around each measuring point. The spectral data of a certain range of surface (controlled by 5 spectral measurement points) in the measurement area around the location of the measurement point was obtained, and the rock and mineral spectral data of the measurement point was obtained by means of average processing value. The matching parameter information such as geographical location, rock and ore types, measurement environment, field photos and sample collection of each measurement point was also recorded and registered according to the requirements.

#### 3.2.3 Pixel Spectral Data Collection

Taking the spectral measurement points measured in the field as the pixel spectral reference acquisition points, 6605 pixel spectral data were collected by self-developed auto-software of hyperspectral pixel acquisition and quality inspection, from the surface reflectance images of GF-5 and ZY1-02D hyperspectral satellites in 16 periods of 20 scenes covering the study area from January 2019 to June 2022. The corresponding attribute parameter information was recorded for each spectral curve. The pixel spectrum acquisition is shown in Table 3.

<table>
<thead>
<tr>
<th>Imaging Time</th>
<th>Payload Name</th>
<th>Image Coverage (scene)</th>
<th>Pixels Collected</th>
</tr>
</thead>
<tbody>
<tr>
<td>20190320</td>
<td>GF-5</td>
<td>1</td>
<td>468</td>
</tr>
<tr>
<td>20190827</td>
<td>GF-5</td>
<td>1</td>
<td>468</td>
</tr>
<tr>
<td>20200727</td>
<td>ZY1-02D</td>
<td>1</td>
<td>56</td>
</tr>
<tr>
<td>20200920</td>
<td>ZY1-02D</td>
<td>2</td>
<td>694</td>
</tr>
<tr>
<td>20210109</td>
<td>ZY1-02D</td>
<td>1</td>
<td>252</td>
</tr>
<tr>
<td>20210122</td>
<td>ZY1-02D</td>
<td>1</td>
<td>468</td>
</tr>
<tr>
<td>20210213</td>
<td>ZY1-02D</td>
<td>1</td>
<td>295</td>
</tr>
<tr>
<td>20210310</td>
<td>ZY1-02D</td>
<td>2</td>
<td>634</td>
</tr>
<tr>
<td>20210507</td>
<td>ZY1-02D</td>
<td>2</td>
<td>741</td>
</tr>
<tr>
<td>20210701</td>
<td>ZY1-02D</td>
<td>1</td>
<td>271</td>
</tr>
<tr>
<td>20210802</td>
<td>ZY1-02D</td>
<td>1</td>
<td>47</td>
</tr>
<tr>
<td>20210926</td>
<td>ZY1-02D</td>
<td>2</td>
<td>936</td>
</tr>
<tr>
<td>20211025</td>
<td>ZY1-02D</td>
<td>1</td>
<td>468</td>
</tr>
<tr>
<td>20211219</td>
<td>ZY1-02D</td>
<td>1</td>
<td>165</td>
</tr>
<tr>
<td>20220114</td>
<td>ZY1-02D</td>
<td>1</td>
<td>174</td>
</tr>
<tr>
<td>20220209</td>
<td>ZY1-02D</td>
<td>1</td>
<td>468</td>
</tr>
</tbody>
</table>

**Table 3. Summary table of Spectrum acquisition period**

### 4. Dataset

#### 4.1 Introduction on Dataset

The hyperspectral surface reflectance dataset (HMTSHR-1) of the typical ore concentration area of Hami Remote Sensing Test Field includes: (1) Surface reflectance images of hyperspectral satellite with a spatial resolution of 30 meters; (2) multi-period pixel spectral data extracted from surface reflectance images of hyperspectral satellites; (3) Spectral data of typical rock and ore measured by ASD in the field; (4) matching parameter data.
<table>
<thead>
<tr>
<th>Data name</th>
<th>Data Quantity</th>
<th>Type</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperspectral Reflectance data</td>
<td>8 periods 8G</td>
<td>Standard remote sensing image format (*.tif)</td>
<td>GF5 or ZY1-02D data of 2019-2022, with 2 period per year All hyperspectral data passing through study areas were collected during Jan. 2019 to June 2022 Geological rock and mineral types and indicators of satellite data performance Measured rock and mineral spectra by ASD in 10:00-17:30 Measurement environment, target type, measurement time, etc.</td>
</tr>
<tr>
<td>Hyperspectral pixel spectral data</td>
<td>6605 bands</td>
<td>ENVI Spectral Library</td>
<td>All hyperspectral data passing through study areas were collected during Jan. 2019 to June 2022 Geological rock and mineral types and indicators of satellite data performance Measured rock and mineral spectra by ASD in 10:00-17:30 Measurement environment, target type, measurement time, etc.</td>
</tr>
<tr>
<td>Parameter description for spectral data of measured rock and ore</td>
<td>20 copies</td>
<td>Excel format</td>
<td>-</td>
</tr>
<tr>
<td>Spectral data of rock and ore measured in the field</td>
<td>468 bands</td>
<td>ENVI Spectral Library</td>
<td>All hyperspectral data passing through study areas were collected during Jan. 2019 to June 2022 Geological rock and mineral types and indicators of satellite data performance Measured rock and mineral spectra by ASD in 10:00-17:30 Measurement environment, target type, measurement time, etc.</td>
</tr>
<tr>
<td>Parameters description for spectral data of measured rock and ore</td>
<td>1 copy</td>
<td>Excel</td>
<td>All hyperspectral data passing through study areas were collected during Jan. 2019 to June 2022 Geological rock and mineral types and indicators of satellite data performance Measured rock and mineral spectra by ASD in 10:00-17:30 Measurement environment, target type, measurement time, etc.</td>
</tr>
</tbody>
</table>

Table 4. Hyperspectral surface reflectance data set of typical ore concentration area in Hami remote sensing test field

4.2 HMTSHR-1 dataset Description

4.2.1 Composition: Figures must be placed in the appropriate location in the document, as close as practical to the reference of the figure in the text. While figures and tables are usually aligned horizontally on the page, large figures and tables can be rotated by 90 degrees. If so, make sure that the top is always on the left-hand side of the page.
(1) Hyperspectral Satellite Surface Reflectance Image Hyperspectral satellite surface reflectance images are stored in GeoTIFF format, containing latitude and longitude coordinate information, and the file name is in the form of “dataset name _ satellite name _ payload name _ imaging date _ dataset name. File suffix”.
(2) Pixel spectral data Regarding to the pixel spectral data, the reflectance values is stored in the spectral general standard format, and the ground spatial information corresponding to each spectral collection point and its ground object attribute information, the load index parameters corresponding to the image data are all stored in excel format. Each of collected spectrum is named in the form of “data set name _ measurement point number _ spectrum collection date _ batch. File suffix”.

4.2.2 Spatial, Temporal and Spectral Characteristics of the Dataset:
(1) Geological Element Unit and Distribution The dataset contains different geological elements such as strata, rock mass, fault tectonic zones, etc. The exposed area and spatial distribution characteristics are different. The distribution of each geological element unit and its corresponding measured spectral points as well as pixel spectral collection points is shown in Figure 6.

Figure 5. Sample display of hyperspectral reflectance data and spectrum of different rock types in study area

Figure 6. Spatial distribution of geological element unit and measure points in study area
(2) Time Distribution
The imaging time of the hyperspectral satellite surface reflectance image is from January 2019 to June 2022, and the time distribution is shown in Figure 7.

![Figure 7. Imaging time of Hyperspectral image](image)

The measurement time of rock and mineral spectral data measured in the field is from 10:00 to 17:30 every day during August to November in 2010, 2011 and 2012, and the time distribution is shown in Figure 8.

![Figure 8. Field measurement time of rock and mineral spectrum](image)

(3) Spectral Type and Scale
The spectral range covers the short-wave infrared band of visible light, and its spectral resolution includes 3nm/10nm, 5nm/10nm, 10nm/20nm. The spectral curve characteristics of typical ground objects at different payload scales are shown in Figure 9.

![Figure 9. Spectrum characteristics of typical ground objects at different payload scale](image)

4.3 Data Quality Control and Accuracy Evaluation
In accordance with relevant industry technical standards and technical specifications of test field construction, to conduct quality control of data products at all levels to ensure the quality and accuracy of data products.

Firstly, the spectral data of rock and mineral measured in the field in the study area are evaluated objectively from waveform characteristics and general trend of spectral curves. Secondly, the accuracy of hyperspectral surface reflectance data is evaluated based on the rock and mineral spectral data measured in the field. The evaluation indexes used in this paper are Root Mean Square Error (RMSE), Correlation Coefficient R and Spectral Angle α (SA).

The mean square error reflects the degree of dispersion between the observed value and the reference truth value, and it is calculated as follows:

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (R_m(i) - R_c(i))^2}
\]  

The correlation coefficient ranges from 0 to 1, reflecting the degree of linear relationship between the observed value and the reference truth value. The closer the correlation coefficient is to 1, the better the consistency is. The formula is as follows:

\[
CC = \frac{\sum_{i=1}^{N}(R_m(i)-\bar{R}_m)(R_c(i)-\bar{R}_c)}{\sqrt{\sum_{i=1}^{N}(R_m(i)-\bar{R}_m)^2 \sum_{i=1}^{N}(R_c(i)-\bar{R}_c)^2}}
\]  

SAM regards the spectrum as a multidimensional vector and evaluates the similarity between the two spectra by calculating the Angle between the pixel and end element spectra in the hyperspectral image. The formula for calculating the spectral Angle is as follows:

\[
\alpha = \cos^{-1} \left( \frac{\sum_{i=1}^{N} R_c(i) R_m(i)}{\sqrt{\sum_{i=1}^{N} R_c(i)^2} \sqrt{\sum_{i=1}^{N} R_m(i)^2}} \right)
\]  

where 

- \( N \) = number of bands of the hyperspectral image 
- \( R_c(i) \) = the reflectance of the reference spectrum in the \( i \) band 
- \( R_m(i) \) = the observation spectrum in the \( i \) band
4.3.1 Quality Evaluation of Field Measured Spectral Data:
According to the apparent color of rock, the ASD spectral instrument parameters were adjusted to ensure the acquisition of high-quality spectral data during field measurement, mainly including the uniform frequency of spectral measurement, dark current measurement and whiteboard measurement.

The field measured rock and mineral spectral data of the same rock type in the study area are overlaid and compared. As shown in Figure 11, the spectral curves of rock and mineral are different in reflectivity, and the overall waveform characteristics of each district and county, especially the diagnostic absorption characteristics of major altered minerals, are significantly consistent.

4.3.2 Accuracy Evaluation of Pixel Spectral Data: In addition to the factors affected by human activities such as mining activities and wind power, the change process of rock ore is relatively slow. During the 10 year cycle, the material composition and internal structure of the rock and ore change slowly, and the spectral characteristics of the rock and ore spectral data are relatively stable. Therefore, the historical data is highly reliable and can be used as a reference data source.
Takng the historical data of rock and mineral spectra measured in the field as reference data, the accuracy evaluation of multi-period pixel spectra was carried out. The accuracy evaluation results of single period were shown in Figure 12, and the accuracy evaluation results of each period were shown in Table 5.

In the accuracy evaluation of single-period surface reflectance data, the mean value of correlation coefficient is 0.921, the mean value of root-mean-square error is 0.087, and the mean value of spectral Angle is 0.074°. In the accuracy evaluation of surface reflectance data of each period, the correlation coefficient of each period is between 0.868 and 0.957, the average correlation coefficient of all periods is 0.911, the root-mean-square error of all periods is between 0.0700 and 0.139, and the average root-mean-square error of all periods is 0.106. The spectral Angle is between 0.046 and 0.130°, and the average spectral Angle for all periods is 0.079°. By comparing the inversion reflectance data of hyperspectral image with the field measured surface reflectance data, the results show that the hyperspectral surface reflectance data has high accuracy.
5. Conclusion

As an important common product in quantitative remote sensing applications, surface reflectance data is the basic data source for most parametric remote sensing products, and it can widely support quantitative applications in natural resources management, ecological environment monitoring as well as urban environment monitoring and other business. Based on the domestic hyperspectral satellite data and the field measured rock and mineral spectral data and spectral measurement points, this paper completed the research on the construction of hyperspectral surface reflectance dataset in the typical ore concentration area of Hami remote sensing test field. The dataset includes GF-5 and ZY1-02D hyperspectral satellite images with 30-meter spatial resolution from 2019 to 2022, multi-period sub-pixel spectral data extracted from hyperspectral satellite surface reflectance images, regional typical field measured rock and mineral spectra and matching parameter data, featuring multi-time series, multi-scale and high spectral resolution. Based on the field measured spectrum data and the matching attribute parameter data, the accuracy of the hyperspectral satellite surface reflectance data is evaluated. The results show that the accuracy of the surface reflectance dataset is good, and the overall correlation coefficient is better than 88%. The hyperspectral multi-scale surface reflectance dataset in this paper combines rich and detailed spectral information with fine spatial information and it can provide good data support for the precise study of spectral characteristics of ground objects (Xiao Xin et al., 2019), simulation and classification of remote sensing application scene (Liang Shuneng et al., 2012; Liang Shuneng et al., 2014), multi-source, multi-feature and well-matched method of altered minerals mapping. Besides, the dataset can also serve the payload design, simulation and demonstration of hyperspectral remote sensing applications on different platforms.

References


