Hyperspectral image classifier based on beach spectral feature

Zhang Liang¹, Gao Lianru², Zhang Bing³

¹ Naval Institute of Hydrographic Surveying and Charting, Tianjin, China
² Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China
³ zltj2004@hotmail.com

Abstract. The seashore, especially coral bank, is sensitive to human activities and environmental changes. A multispectral image, with coarse spectral resolution, is inadaptable for identifying subtle spectral distinctions between various beaches. To the contrary, hyperspectral image with narrow and consecutive channels increases our capability to retrieve minor spectral features which is suitable for identification and classification of surface materials on the shore. Herein, this paper used airborne hyperspectral data, in addition to ground spectral data to study the beaches in Qingdao. The image data first went through image pretreatment to deal with the disturbance of noise, radiation inconsistence and distortion. In succession, the reflection spectrum, the derivative spectrum and the spectral absorption features of the beach surface were inspected in search of diagnostic features. Hence, spectra indices specific for the unique environment of seashore were developed. According to expert decisions based on image spectrums, the beaches are finally classified into sand beach, rock beach, vegetation beach, mud beach, bare land and water. In situ surveying reflection spectrum from GER1500 field spectrometer validated the classification production. In conclusion, the classification approach under expert decision based on feature spectrum is proved to be feasible for beaches.

1. Introduction
Some islands of China are occupied by other countries and one third of Chinese territorial waters are still in dispute. Beach, as the interaction zone of land and sea, is sensitive to and indicative of global environmental changes [1]. Global warming and human activities bring impressive changes to beaches. For example, the coral bank of lots of tropic islands is sinking along with the degraded coral reef and increasing sea level [2]. Thorough investigation of beach is critical for maintaining Chinese rights and interests of ocean. Conventional in-situ survey is hard to perform either on foot or on boat, since the sea shore is dynamically dry or flood with the tide. In the meantime, multispectral remote sensing is not suitable for fine classification of sea shore due to its coarse spectral resolution. Therefore, an effective technology is needed to monitor the variable beaches, especially the island bank. Hyperspectral remote sensing provides a feasible choice for us to fulfill the performance. Mild differences between diverse beaches can be identified with the retrieved continuous image spectrum. However, traditional statistical classification methods do not fit for classification of hyperspectral images, as these methods are generally difficult to discriminate the beaches with similar reflectance, and it is difficult to find adequate samples for approving accuracy.

Based on painstaking investigation of spectrum characteristic of the variable beaches, an operational non-parametric classifier was proposed for beach classification, which constructed the
decision tree using spectral indices derived from feature spectrum. Another important advantage of the proposed approach is no assumption about distributions of data is need.

2. Data collection and preprocessing

2.1. Data Collection
A hyperspectral experiment was carried out on the seashore of Xuejia Island in Qingdao on September 8 in 2002. Consistent airborne data, satellite image and ground spectrum were collected for comparison. The target site composed of four kinds of beaches, which are the rock beach, the sand beach, the vegetable beach and the mud beach.

The airborne hyperspectral image was gathered from 10:00 to 16:00, which parallels to the bank, with OMIS imaging spectrometer made in the Shanghai Institute of Technical Physics, China. The flight maintains constant speed of 150 kilometers per hour and invariable height of 800 meters. The ground resolution of the hyperspectral image was 2.4 meters, recorded in 128 bands and 12 bits.

The PRISM satellite image covering the same region was obtained with ground resolution of 2.5 meters, which will be used for geometrical correction of the airborne hyperspectral image.

Simultaneous ground spectrum was acquired with GER 1500 spectrometer, referred to the relevant official standards. The spectrums were recorded in 512 bands ranging from 276.8nm to 1097.3nm. The center of the relative uniform beach is carefully selected as the representatives. Instrument optimization was performed before each spectrum collection to assure consistency of spectral response. The spectrums were collected with a three-FOV optic, twelve to eighteen centimeters right above the site. Each spectrum measurement was fulfilled in one minute to minimize irradiation variation. A calibrated white reference panel was inspected before and after each target measurement. As a result, more than one thousand pieces of spectrum data were collected from the test site. In addition to the four types of beaches, the spectrums of the asphalt road, cement road, water and constructions were gathered.

2.2. Data preprocessing
Data Preprocessing is necessary for the three kinds of experimental data. Precise geometrical rectification was first performed to the PRISM satellite image. Then, the airborne hyperspectral images were rectified one by one, through image registration with the same PRISM image. Before the geometrical rectification, relative and absolute radiation correction was performed to eliminate radiation error. The relative calibration causes the measurements internally consistent with each other; whereas absolute calibration makes the data in agreement with an external standard to get apparent radiance or apparent reflectance [3]. The atmospheric correction should also be carried out in precedence to retrieve image reflectance in conformity to the ground collection, which is done by empirical linear relationship method to convert data from DN to reflectance.

Furthermore, for the ground spectrums, we also perform some steps before they can be used. First, the sun zenith angle was got from the GPS data. Then, the bi-directional reflectance characteristic of the reference panel can be retrieved and the reflectance curve of the target can subsequently deduced from the bi-directional reflectance factor. Finally, average of each group of spectrum waves were used for representation of appropriate sample site.

Then spectrum equivalent calculation is performed for the GER and OMIS data before comparison. Similar targets may present distinct spectrum, since different spectral resolution leads to different mean radiance intensity recorded by the sensors [4].

3. Beach classification based on feature spectrum

3.1. Spectrum analysis
Detailed spectrum analysis is necessary for precise classification of hyperspectral image. As can be seen from figure 1, some discipline can be concluded from the field reflectance spectrum. In general,
the reflectance of road or manmade buildings is always higher than all kinds of beaches in the visible to near-infrared region. On the contrary, the reflectance of seawater is all along lower than all kinds of beaches in the same region. Dry sand beach and wet sand beach have similar reflectance in the blue to green region, while the wet sand beach gives much fainter reflectance than the dry one in the red to near-infrared region. The reflectance curves of vegetation beach and rock beach have resembling red vale shape, however, the reflectance of vegetation is quite stronger than rock in the near-infrared region. The reflectance of mud beach keeps the lowest among all the beaches in the blue to red region. Image spectrum should also be considered when we classify hyperspectral image of seashore. There are many special phenomena in this area, such as whitecap and glint. The image spectrum of whitecap has three apices, as shown in figure 2. Based on the above analysis, we can draw a conclusion and develop appropriate indices to discriminate the coastal targets [5].

Figure 1. Ground reflectance spectrum of coastal targets

Figure 2. Image reflectance spectrum of whitecap
3.2. Beach classification

The spectrum of the same kind of beach may be different in various conditions, such as different humidity or irradiation environment, which is related with the time of image acquirement. Moreover, variable beaches may be inseparable due to the mild spectrum difference between each other, in contrary to the significant differentiation between beach and land objects. In order to discriminate all kinds of targets in the seashore image, a decision tree classification approach was developed, based on specified spectrum indices. The decision tree was designed to divide the image step by step, from the most extraordinary spectrum to the most similar group. More significant different is the spectrum to be, higher priority would be granted to it for segmentation. The proposed approach is present as follows with several aspects.

First, the cloud is the most significant object with regard to all other objects on the ground, which has extraordinary bright spectrum in visible range. Hence, a cloud index was derived from the mean of reflectance in visible region, such that the cloud was retrieved.

Second, pixels of water were always abundant in coastal image with unique descending spectrum curves in near-infrared region. Therefore, the classification procedure can be speeded up by removing a large amount of water pixels. The corresponding spectral index for water was defined using the odds between the reflectance of 843.5 nm and that of 675.8 nm.

Whitecap is the foam of the wave, a frequent phenomenon in coast shore image, whose distinct spectrum has three apices. A set of indices are built upon this feature spectrum. The whitecap pixels are subsequently classified.

\[
FSI\_NIR = R_{767.1} - 2 \cdot R_{626.5} + R_{540.5}
\]

\[
FSI\_Red = R_{626.5} - 2 \cdot R_{540.5} + R_{455.7}
\]

\[
FSI\_Green = R_{540.5} - 2 \cdot R_{455.7} + R_{353.7}
\]

Thereafter, types of beaches were segmented into two groups through vegetable index, one group includes the rock beach and the vegetable with higher vegetable index, another group is comprised of the sand beach, bare land and the mud beach with lower values for vegetable index. Then, the sand beach and the bare land can be separated from the mud beach, since their reflectance on the green spectrum is much higher than the mud beach.

Rock beach was the next targets to be retrieved, whose spectrum has similar red edge as the vegetable, since a lot of rocks on the beach were almost sheltered by micro algae, coral, moss or grass. The most distinct difference between the spectrum of vegetable and rocks consists in near-infrared region. The mean of reflectance in this region was used to construct another index, which is higher for vegetable and lower for rock. The index can also be used to discriminate sand beach and bare land. Sand beach has higher value for this index than the bare land.

Until now, the associate beaches and objects were all classified successfully and an operational beach classification approach is developed.

4. Classification results

Sixty six bands of the OMIS hyperspectral image covering the visible Near-infrared region, from 455.7 nanometers to 1097 nanometers, were selected for beach classification in the experiments. Figure 3 is part of the false color image for testing, where red, green and blue channel was derived from 810.6nm, 626.5nm and 553.1nm wavelength bands respectively. The hyperspectral image was segmented into six classes using the decision tree approach based on feature spectrum. Four main kinds of beaches were set apart precisely, validated by field survey. It is worth noting that the sand beach and bare land, the rock beach and vegetation, which are targets with similar spectrum figure, were successfully classified, as we can see from figure 4.
Several Region Of Interest (ROI) of target classes were randomly selected and recorded, in order to quantitatively evaluate the outcome accuracy. A pair of data was chosen from each pixel of the ROI.
One is the results of classification; the other is the true attribute according to field survey and investigation. Then, statistics was taken to get confusion matrix and precision of classification. Table 1 presents the obtained confusion matrix, while Table 2 revealed the precision of classification. In total, 15635 pixels in different areas were collected. Kappa coefficient calculated from the confusion matrix was 0.8653, preceding traditional result. The total classification precision was 89.7474%, which further validated the accuracy of the classification approach of decision tree based on feature spectrum.

Table 1. Confusion Matrix of coastal zone classification

| Class       | vegetation | bare land | sand beach | rock beach | mud beach | seawater | Total |
|-------------|------------|-----------|------------|------------|-----------|----------|-------|
| vegetation  | 1041       | 145       | 0          | 15         | 1         | 0        | 1202  |
| bare land   | 0          | 1841      | 331        | 0          | 50        | 1        | 2223  |
| sand beach  | 0          | 314       | 5398       | 0          | 46        | 2        | 5760  |
| rock beach  | 0          | 510       | 0          | 1953       | 113       | 0        | 2576  |
| mud beach   | 0          | 0         | 0          | 0          | 282       | 13       | 295   |
| seawater    | 0          | 4         | 0          | 0          | 58        | 3517     | 3579  |
| Total       | 1041       | 2814      | 5729       | 1968       | 550       | 3533     | 15635 |

Table 2. Precision of coastal zone classification

| Class       | Producer Precision | User Precision | Producer Precision | User Precision |
|-------------|--------------------|----------------|--------------------|----------------|
| Vegetation  | 100.00             | 86.61          | 1041/1041          | 1041/1202      |
| bare land   | 65.42              | 82.82          | 1841/2814          | 1841/2223      |
| sand beach  | 94.22              | 93.72          | 5398/5729          | 5398/5760      |
| rock beach  | 99.24              | 75.82          | 1953/1968          | 1953/2576      |
| mud beach   | 51.27              | 95.59          | 282/3533           | 282/295        |
| Seawater    | 99.55              | 98.27          | 3533/3579          | 3517/3579      |

5. Conclusions
In this work, the decision tree classification method grounded on feature spectrum was evaluated, using a scene of coastal hyperspectral image. The proposed approach is simple but robust, hence, it is adapted to be programmed and be put into coastal operational applications. More abundant expert knowledge can also be conveniently imported and amended according to practical applications. Hybrid decision tree classification method can be developed as well, by introducing other efficient methods. This type of approaches would be applicable in extensive coastal engineering and investigation.

6. Acknowledgement
This work was supported by National Key Technology R&D Program of China under Grant 2011BAH12B04, and by National Natural Science Foundation of China under Grant 40901225.

References
[1] Zhang L, Zhang B, Chen Z C, Zheng L F and Tong Q X 2009 Acta Oceanologica Sinica 28 1-13
[2] Zhang L, Mao Z H, Zhang B, Zheng L F and Tong Q X 2008 Developing operational approach for beach classification using hyperspectral data: Proc. of Information Technology and Environmental System Sciences (Jiaozuo, China, 15-17 May 2008) 2, 267-272
[3] David A L 2003 Signal Theory Methods in Multispectral Remote Sensing (New Jersey, John Wiley & Sons) chapt 5 pp 270-271
[4] Mao Z H 2003 The comparison between the in-situ measured and the air-boarding hyperspectral measured radiance (*Second Institute of Oceanography, State Oceanic Administration Report*) p 5

[5] Barducci A, Donatella G, Paolo M and Ivan P 2009 *Journal of Environmental Management* **90** 2278-2286