

# Detection and Classification of Grassland with *Stellera chamaejasme* through HJ-1A Hyperspectral Image in Northern Tibet

Fenfen Guo<sup>1,a</sup>, Jingjing Xiao<sup>2,b</sup>, Jianrong Fan<sup>3,c</sup>

<sup>1</sup> Second Institute of Oceanography, SOA, Hangzhou, Zhejiang, 310012, China

<sup>2</sup> Zhejiang Climate Center, Hangzhou, Zhejiang, 310012, China

<sup>3</sup> Institute of Mountain Hazards and Environment, Chinese Academy of Sciences & Ministry of Water Conservancy, Chengdu, Sichuan, 610041, China

<sup>a</sup> [gufefine@163.com](mailto:gufefine@163.com), <sup>b</sup> [xiaojingjing2005@163.com](mailto:xiaojingjing2005@163.com), <sup>c</sup> [fjr@imde.ac.cn](mailto:fjr@imde.ac.cn)

**Keywords:** Northern Tibet, *Stellera chamaejasme*., Grassland degradation, HJ-1A, Hyperspectral.

**Abstract.** *Stellera chamaejasme* is the most typical invasive plant in North Tibet. Monitoring the distribution of the *Stellera chamaejasme* is the prerequisite for grassland quality evaluation or management. In this paper, the hyperspectral image of HJ-1A satellite (HJ-1A HSI) was used to monitor the *Stellera chamaejasme* in the northern region of Xainza. Linear Spectral Unmixing (LSU) algorithm was used to get the abundances of six landuse types(include grassland with *Stellera chamaejasme*) based on HSI. Feature extraction method of principle component analysis (PCA) was used to compress the HSI information into five bands. And then, maximum likelihood classifier(MLC) was applied to the six abundances and the first five principal components for getting the classification result. The result showed that the classification result of LSU-MLC was better than PCA5-MLC, the mapping accuracy of *Stellera chamaejasme* grassland based on both methods were all greater than 91%, the user accuracy was all greater than 66%. We concluded that HJ-1A HSI can be used to detect and classify *Stellera chamaejasme* in northern Tibet. In addition, we got the spatial distribution characteristics of *Stellera chamaejasme* by overlay analysis of classification result with the terrain factors: the *Stellera chamaejasme* mainly distributed in the area with elevation between 4500 and 4800m, slope less than 10° and mostly in the Southwest aspect, where radiation and hydrothermal condition were favorable.

## Introduction

*Stellera chamaejasme* in northern Tibet is the indicator of grassland degradation [1]. *Stellera chamaejasme* hampers other plants growth, and high intake would make cattle and sheep syncope [2]. In this context, detecting and mapping the distribution of *Stellera chamaejasme* across large areas is crucial to effective land management. The landscape of grassland with *Stellera chamaejasme* is similar with the bad growing grassland. As one development trend of the remote sensing technology, hyperspectral remote sensing can crack different objects with same spectrum, and identify similar types of grassland by high spectral resolution at regional scale [3,4]. The vacancy of hyperspectral image (HSI) in China is made up by HJ-1A satellite. We did this study in order to get the applicability of HJ-1A HSI for similar landscapes classification.

Through the research of domestic and foreign scholars, the classification methods based on statistical features and spectral characters work well in HSI classification [4,5]. The prerequisite for the former is feature extraction [6]; the latter can be divided into two types: spectral matching and mixed pixel decomposition. Mixed pixel decomposition is used to get the abundance value of each material [7]. As a way of mixed pixel decomposition, linear spectral mixture model is widely used at home and abroad [8].

In this paper, HJ-1A HSI was used to monitor the *Stellera chamaejasme* in the northern region of Xainza. Linear Spectral Unmixing (LSU) algorithm based on the spectral characters was used to get the abundances of six landuse types (include *Stellera chamaejasme* grassland) based on HSI. The feature extraction method of PCA was used to compress the HSI information into five bands.

Following data processing, the popular supervised maximum likelihood classifier (MLC) was used to the six abundances and the first five principal components for getting the *Stellera chamaejasme* distribution area. Furthermore, this paper analysed *Stellera chamaejasme* distribution characteristic based on classification result and terrain factors.

### Study site and dataset

**Study Site.** The study area was located in the middle-south of the Qiangtang Plateau of northern Tibet, which situated between the Jiagang mountain and the Selin Lake (88 ° 34'23 "E—89 ° 14'14" E, 30°45'44 "N—31°32'44" N). The elevation ranges was from 4720 m to 6450m as derived from 30 m resolution digital elevation model (DEM) data. The region geomorphic type was the typical plateau lake-basin landscape. This area belonged to the typical plateau continental climate, with the characteristics of low temperature, large temperature difference between day and night, few and concentrated rainfall, et al.

The main landuse types in the study area were grassland, water and bare land. Based on classification approach of vegetation genesis, the grassland in the study area had tress subclasses, swamp meadow, alpine meadow and alpine steppe [9]. The vegetation coverage of swamp meadow and alpine meadow was greater than 65%. The vegetation coverage of alpine steppe was in the range of 5~50%. According to the level of vegetation coverage, the alpine steppe was subdivided into med-cover alpine steppe (20 ~ 50%) and low-cover alpine steppe (5 ~ 20%) in this paper. In recent years, *Stellera chamaejasme* was widespread in the alpine steppe due to grassland degradation. The landscape of *Stellera chamaejasme* steppe was similar with low-cover alpine steppe.

**Field Data Collection.** We collected field data of six types (swamp meadow, alpine meadow, med-cover alpine steppe, low-cover alpine steppe, *Stellera chamaejasme* steppe and bare land ). A total of 120 patches (20 patches for each type) were randomly chosen in the study area and traced using the GPS equipment. All patches were imported into ENVI software and overlaid onto the HJ-1A HSI. The pixels of HSI covered by these patches were used for selecting training samples and verifying results.

**Remote Sensing Images.** In this paper, we used the HJ-1A (One satellite of environment disaster monitors constellation) HSI acquired from China Center for Resources Satellite Data and Application (<http://www.cresda.com>). The images with 3% cloud were acquired on 30 June 2012, of which the spatial resolution is 100m and the spectral range is 459 ~ 956nm and the spectral resolution is less than 10nm, has a total of 115 bands. We removed the first 10 bands with obvious vertical stripes which would influence the interpolation process and the final classification result seriously. The spectral range of the remaining image was 481 ~ 956nm.

**Auxiliary Data.** The 30m-resolution STRM DEM data and the TM image data (GLS2005) on October 3, 2007 retrieved from the International Scientific Data Sharing Service Platform (<http://datamirror.csdb.cn>) were also analyzed as the auxiliary data.

### Image pre-processing

**Image pre-processing.** Before image information extraction, the HSI needed to be made a series of pre-processing, including five steps: format conversion, atmosphere correction, image mosaic, study area image clipping and geometric correction. Firstly, the format of HSI was converted to geotiff format by the h5 format conversion program and the value was divided by scaling factor 100. Secondly, the FLASHH model was used to atmosphere correction of HSI [10]. In addition, study area across two images, images mosaic and clipping were operated according to the data range of study area. The image geometric correction was based on the TM images. Image pre-processing were accomplished by ENVI software.

**Background masking.** We focused only on the six landuse types by creating a mask using preliminary classification. Three types including cloud, shadow and water were identified by supervised

classification, and merged to build mask layer in data preprocessing. This mask was applied to all subsequent image processing methods in this study.

## Classification Methodology

**Principle Component Analysis.** Feature extraction method plays important roles in hyperspectral image processing [11]. Principle component analysis (PCA) is considered as the most important data transformation method to keep the most useful information with the least number of bands or dimensions in the HSI [12]. In this paper, we specifically examined the feature extraction methods of PCA. It linearly transformed the data into a subspace while retaining the greatest variance. The transformation was based on the covariance or correlation matrix of the dataset, it worked well in remote sensing because the variances of classes were frequently distributed in the direction of maximum data scatter.

The eigenvalue and accumulated information of the first five principal components extracted from HSI image were listed in table 1. It showed that the accumulated information of the first five principal components area accounts for 98.66 percent of the total information, the first principal component area accounts for 91.61percent.

Table 1 The eigenvalue and accumulated information of the first five principal components extracted from HSI image

| PC | eigenvalue | accumulated information (%) |
|----|------------|-----------------------------|
| 1  | 144508.13  | 91.16                       |
| 2  | 8989.36    | 96.83                       |
| 3  | 2122.81    | 98.17                       |
| 4  | 406.40     | 98.43                       |
| 5  | 373.29     | 98.66                       |

**Linear Spectral Unmixing.** Linear Spectral Unmixing (LSU) solves for the abundance values of each material for every pixel. It's the representative HSI classification method based on spectral feature. The theory is that the reflectance at each pixel of the image is assumed to be a linear combination of the reflectance of each material present within the pixel. The theory equation can be established as follow:

$$\rho_i = \sum_{j=1}^n \alpha_{ij} f_j + \varepsilon_i \quad (1)$$

Where  $\rho_i$  is the reflectance of any i-band in the image, n is the number of the materials,  $\alpha_{ij}$  is the reflectance of the j- material in the i-th band, and  $f_j$  is the proportion of the material in the pixel. The model is constrained by two factors. One is the sum of the proportion of the material is 1; the other was that the proportion of all materials are non-negative. Linear Spectral Unmixing calculates the abundance value of each material based on the least square method.

We got the abundance value of the six landuse types in the study area based on LSU and HJ-1A HSI. There was striping in each abundance value image. And the abundance value of *Stellera chamaejasme* steppe was not outstanding over each other. For this reason, it is impossible to get the distribution area of *Stellera chamaejasme* steppe quite precisely by threshold segmentation.

**Maximum Likelihood Classifier.** The maximum likelihood classifier (MLC) is one of the most broadly used supervised classification methods in multispectral remote sensing. A pixel with the maximum posterior probability is categorized into the corresponding class [13]. The features and abundances extracted from HSI were similar to the multispectral image with few bands. The MLC and field sample data were used to the above data to get the classification results. It avoided segmenting threshold subjectively for abundances data, and taked full advantage of field sample data.

**Accuracy assessment.** Accuracy assessment is an important link in the remote sensing technology application. Confusion matrix is the basic index to evaluate the classification accuracy [13]. It is a comparison array for expressing the pixels ratio of classification to ground survey. We can get the pixel

number of correctly or wrongly classified for each classification type from confusion matrix, But it can't show the good and bad points of classification accuracy intuitively. So the indexes of accuracy assessment derived from confusion matrix were used in the study to quantify the classification accuracy for each classifier, which include: overall accuracy, kappa coefficient [14], user accuracy and producer accuracy.

## Results and discussions

**Classification Result.** The MLC were applied to features of PCA and abundances data of LSU. We got the classification maps of both methods. Each map had six landuse types (swamp meadow, alpine meadow, med -cover alpine steppe, low-cover alpine steppe, Stellera chamaejasme steppe and bare land). All classification results were listed in fig. 4. Blank area in the maps was the mask of cloud, shadow and water. Confusion matrix derived from PCA5-MLC classifier was listed in table 2. Confusion matrix derived from LSU-MLC classifier was listed in table 3.

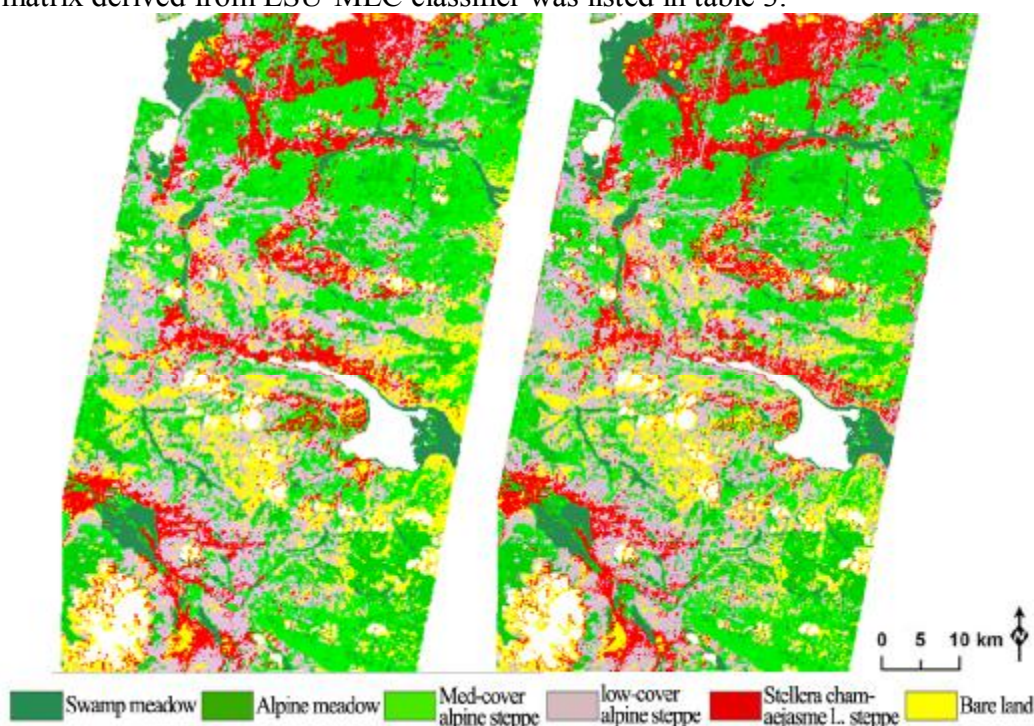


Fig.1 Classification map generated from the PCA5-MLC classifier (left) and LSU-MLC classifier (right). The black area indicated the mask pixels.

Table 2 Confusion matrix derived from PCA5-MLC classifier. The overall accuracy was 80.53% and the kappa coefficient was 0.76.



| Classification types        | Ground truth classes (%) |               |                          |                         |                             |           | Mapping accuracy | User accuracy |
|-----------------------------|--------------------------|---------------|--------------------------|-------------------------|-----------------------------|-----------|------------------|---------------|
|                             | Swamp meadow             | Alpine meadow | Med -cover alpine steppe | Low-cover alpine steppe | Stellera chamaejasme steppe | Bare land |                  |               |
| Swamp meadow                | 98.44                    | 0             | 0                        | 0                       | 0                           | 0         | 98.44            | 100.00        |
| Alpine meadow               | 1.56                     | 77.73         | 0.68                     | 0                       | 0                           | 0         | 77.73            | 97.62         |
| Med-cover alpine steppe     | 0                        | 9.00          | 83.56                    | 0.61                    | 0                           | 0         | 83.56            | 85.92         |
| low-cover alpine steppe     | 0                        | 10.43         | 15.07                    | 50.92                   | 3.39                        | 8.57      | 50.92            | 59.29         |
| Stellera chamaejasme steppe | 0                        | 0             | 0                        | 28.22                   | 91.53                       | 8.57      | 91.53            | 66.26         |
| Bare land                   | 0                        | 2.84          | 0.68                     | 20.25                   | 5.08                        | 82.86     | 82.86            | 65.41         |

**Table 3 Confusion matrix derived from LSU-MLC classifier. The overall accuracy was 82.14% and the kappa coefficient was 0.78.**

| Classification types        | Ground truth classes (%) |               |                         |                         |                             |           | Mapping accuracy | User accuracy |
|-----------------------------|--------------------------|---------------|-------------------------|-------------------------|-----------------------------|-----------|------------------|---------------|
|                             | Swamp meadow             | Alpine meadow | Med-cover alpine steppe | low-cover alpine steppe | Stellera chamaejasme steppe | Bare land |                  |               |
| Swamp meadow                | 97.92                    | 0             | 0                       | 0                       | 0                           | 0         | 97.92            | 100           |
| Alpine meadow               | 1.04                     | 73.46         | 0                       | 0                       | 0                           | 0         | 73.46            | 98.73         |
| Med-cover alpine steppe     | 1.04                     | 15.17         | 83.56                   | 0                       | 0                           | 0         | 83.56            | 78.21         |
| low-cover alpine steppe     | 0                        | 10.43         | 14.38                   | 66.87                   | 3.39                        | 7.62      | 66.87            | 66.46         |
| Stellera chamaejasme steppe | 0                        | 0             | 0                       | 18.4                    | 92.37                       | 11.43     | 92.37            | 72.19         |
| Bare land                   | 0                        | 0.95          | 2.05                    | 14.72                   | 4.24                        | 80.95     | 80.95            | 71.43         |

Fi

g.1 showed that the classification maps generated from PCA5-MLC and LMSS-MLC classifiers were similar. According to calculation results, the two classification maps had correlation coefficient of 0.73. The distribution area of Stellera chamaejasme steppe in both classification maps had correlation coefficient of 0.77. The high relativity and consistency proved that the classification results were reasonable.

Comparative analysis of table 1 and table 2 showed that the PCA5-MLC classifier had relatively lower accuracy than the LSU-MLC classifier. The overall accuracy of PCA-MLC classifier was 80.53% and the kappa coefficient was 0.76. The Stellera chamaejasme steppe mapping accuracy by PCA-MLC classifier was 91.53% and user accuracy was 66.26%. The overall accuracy of LSU-MLC classifier was 82.14% and the kappa coefficient is 0.78. The Stellera chamaejasme steppe mapping accuracy by LSU-MLC classifier was 92.37% and user accuracy was 72.19%. In addition, the plaque distribution of Stellera chamaejasme steppe in classification map was consistency with field survey result, which generated from the LSU-MLC classifier.

The classification error of LSU-MLC classifier was mainly manifested on that Stellera chamaejasme steppe was confused with low-cover alpine steppe or bare land, pixels around the mask area were classified to Stellera chamaejasme steppe falsely. On the other hand, HJ-1A HSI still contained stripe noise after removing the first 10 bands. For this reason, the classification map exhibit north-South stripe.

**Stellera chamaejasme Distribution Characteristics.** Stellera chamaejasme steppe classified by LSU-MLC classifier accounted for 15.51% of the study area. It is mainly distributed around the rivers and lakes, scattered elsewhere. In order to further analyse the distribution characteristic of the Stellera chamaejasme. The overlay analysis was used to terrain factors and Stellera chamaejasme steppe distribution area detected by LSU-MLC classifier.

The left part of the Fig.2 was the distribution histograms of three terrain factors(elevation, slop and aspect) distribution for the whole study area. The elevation of the study area was mainly concentrated in 4550~5000m (Fig.2a), the slope was mainly below 15 degree (Fig.2b), the 8 aspects(the values of 1~8 represented north, northeast, East, Southeast, South, southwest, West and northwest) were relatively homogeneous , northeast aspect was a little more(Fig.2c).

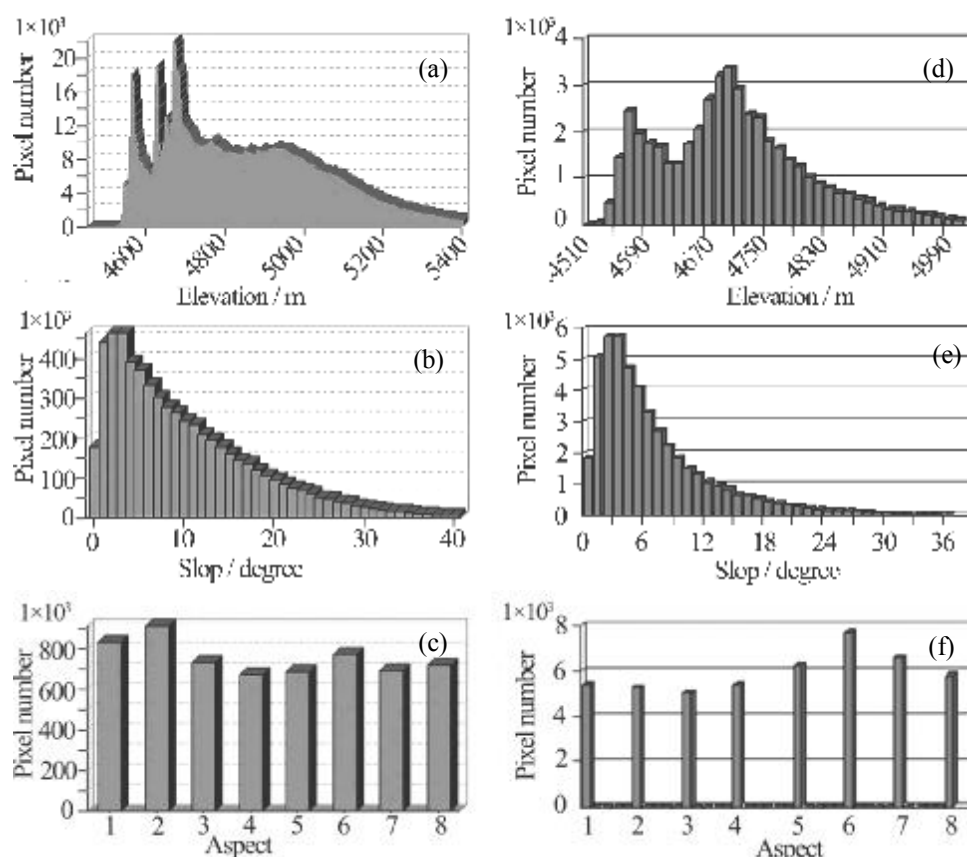


Fig.2 The distribution histograms of three terrain factors (elevation, slop and aspect) for the whole study area (left) and for the area of Stellera chamaejasme steppe (right)

The right part of Fig.2 was the distribution histograms of three terrain factors for the area of Stellera chamaejasme steppe. We got the information as follow from the figure: The Stellera chamaejasme steppe was mainly concentrated in 4500-4800m and two peaks located at 4500m and 4700m. These two peaks were similar to the study area elevation (Fig.2d). The slop of Stellera chamaejasme steppe distribution area was mainly below 10 degree, one peak located at 3, From the beginning of the 5 degree, it gradually reduced with the increase of slope (Fig.2e).The Stellera chamaejasme steppe area in southern were more than that northern, and the west was more than east. In all eight aspects, the Stellera chamaejasme steppe area in southwest was more than any other one (Fig.2f).

In summary, in the study area the Stellera chamaejasme mainly distributed in the area with elevation between 4500 and 4800m, slope less than  $10^\circ$  and mostly in the Southwest aspect, where radiation and hydrothermal condition were favorable.

## Conclusion

In this paper, we got the classification results of PCA5-MLC classifier and LSU-MLC classifier based on HJ - 1 A HSI and field collection data. The classification maps of both classifier had correlation coefficient of 0.73. The mapping accuracy of the *Stellera chamaejasme* steppe based on both methods were all greater than 91%, the user accuracy was all greater than 66%. The high classification accuracy illustrated that HJ-1A HSI can be used to detect and classify the *Stellera chamaejasme* in northern Tibet.

The LSU-MLC classifier with higher classification accuracy was more suitable for *Stellera chamaejasme* detection, took full advantage of spectral information and field sample data, avoided segmenting threshold subjectively. This method could be used on grassland resource investigation, assessment and management in northern Tibet as a cost-effective way.

The analysis of the *Stellera chamaejasme* distribution characteristic showed that *Stellera chamaejasme* in the study area mainly distributed in the area with elevation between 4500 and 4800m, slope less than 10° and mostly in the Southwest aspect, where radiation and hydrothermal condition were favorable. The conclusion was consistent with the characteristics of ecology research conclusions, which proved the reliability of the classification results from the other aspect.

It is worthy to note that the classification map still exhibit north-south stripe although removing the first 10 bands of HJ-1A HSI during the processing. Next we will do more research on image noise removing to improve classification accuracy.

## References

- [1] Y. Liu, R.J. Long and T. Yao: Research Progress on *Stellera Chamaejasme* in Grassland, *Pratacultural Science*. Vol.6(2004) , p.55-61.
- [2] J.B. Qian: Field Measurement and Satellite Remote Sensing Estimation of *Stellera Chamaejasme* in Qilian Mountains, master, Lanzhou University (2009).
- [3] Y. Yang: The Key Diagnosis Technology of Rice Blast Based on Hyper-spectral Image, doctor, Zhejiang University(2012).
- [4] C.I. Chang: Hyperspectral Data Exploitation: Theory and Applications, *Photogrammetric Engineering & Remote Sensing*. Vol.6( 2007) , p.441-442.
- [5] J. Li: Research on Dimensionality Reduction and Classification of Hyperspectral Image, master, Cental South University (2012).
- [6] P. Hsu and Y. Tseng: Feature Extraction for Hyperspectral Image(Proceedings of the 20th Asian Conference on Remote Sensing, Hong Kong 1999).
- [7] L. Parra, C. Spence and P. Sajda: Unmixing Hyperspectral Data, *Advances in Neural Information Processing Systems*. Vol.12(2000) , p. 942-948.
- [8] F.Chen, Q.Y. Qiu and Y.Z. Xiong: Pixel Unmixing Based on Linear Spectral Mixture Model: Methods and Comparison, *Remote Sensing Information*. Vol.4(2010) , p.22-28.
- [9] Q. Liu: The Classification of Natural Grassland in China, *Sichuan Grassland*. Vol. 2(1996), p. 1-5.
- [10] T. Cooley: FLAASH, a MODTRAN4-based Atmospheric Correction Algorithm in: *Geoscience and Remote Sensing Symposium*. Vol 3(2002), p. 1414- 1418.
- [11] E. Underwood, S. Ustin and D. Dipietro: Mapping Nonnative Plants Using Hyperspectral Imagery, *Remote Sensing of Environment*. Vol 86 (2003), p. 150–161.
- [12] J.K. Zhang: Research on Feature Extraction of Remote Sensing Images, master, Northwestern Polytechnical University (2007).
- [13] S.H. Li, J.L. Wang and Y. Bi: A Review of Methods for Classification of Remote Sensing Images, *Remote Sensing for Land & Resources*. Vol.2(2005), p.1-6.
- [14] R. G. Congalton: A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*. Vol.1(1991), p.35-46.