# FEATURE EXTRACTION AND CLASSIFICATION OF HYPERSPECTRAL IMAGES

A.Munkh-Erdene and M.Ganzorig

Institute of Informatics and RS, Mongolian Academy of Sciences
Ave.Enkhtaivan-54B, Ulaanbaatar-51, Mongolia; Tel: 976-11-453660, Fax: 976-11-458090
*E-mail: munkherdene@informatic.ac.mn*
*E-mail: ganzorig@arvis.ac.mn*

**KEY WORDS:** Hyperspectral, feature extraction, classification

## ABSTRACT

In recent years, the processing and analysis of hyperspectral images have become the main tasks of many researchers dealing with RS image processing. Unlike the traditional multispectral datasets taken in the optical range of electro-magnetic spectrum, the hyperspectral data deals with an enormous amount of bands and the data are formed as collections of hundreds of images of the same scene with each image corresponding to a narrow interval of the electro-magnetic wavelength. It is clear that such datasets offer the superior potential for more accurate and detailed information extraction than is possible with other types of RS data.

In this research, Hyperspectral Imager (Hyperion) and Advanced Land Imager (ALI) sensor images have been used onboard EO-1 satellite. The goal of this paper is to compare two different approaches in geological feature extraction for an image classification. Before the classification spectral and spatial enhancements are applied. Advanced satellite images classification represents an accurate and cost effective for land cover mapping at regional scale.The output of each of the feature extraction method is classified using a maximum likelihood classification and spectral angle mapper methods. The results are analyzed and compared.

## 1. INTRODUCTION

At present, the processing and analysis of hyperspectral images have become the main tasks of many researchers dealing with RS image processing. Unlike the traditional multispectral datasets taken in the optical range of electro-magnetic spectrum, the hyperspectral data deals with an enormous amount of bands and the data are formed as collections of hundreds of images of the same scene with each image corresponding to a narrow interval of the electro-magnetic wavelength. It is clear that such datasets offer the superior potential for more accurate and detailed information extraction than is possible with other types of RS data. This means that hyperspectral data sets provide a wealth of information and are used for many different applications such as geological investigation, forest change analysis, environmental mapping, global change study, wetlands mapping, crop analysis, traffic ability assessment, plant identification, mineral recognition, and many others (Plaza *et al.* 2012, Amarsaikhan *et al.* 2011).

As the hyperspectral images consist of a large number of bands, their unique characteristics pose different processing problems, which could be necessarily tackled under specific mathematical formalisms, such as segmentation and classification as well as spectral mixture analysis (Smith *et al.* 1990 and Jia *et al.* 1999). In order to reduce the data volume, the techniques for reducing the image dimensionality are often applied. Usually, it is reduced by applying different transformation techniques by retaining only the significant components for further processing. Information extraction is generally done through classification of the images and identifies which pixels contain a variety of spectrally distinct labels. Many attempts are being made to reduce the data dimensionality and extract reliable information needed for different decision-making (Amarsaikhan and Ganzorig 1999, Keshava and Mustard 2002, Richards 2005, Tsai *et al.* 2009, Yang *et al.* 2009, Vega *et al.* 2012).

## 2. DATA SOURCES AND TEST SITE

In the current study, 242 band Hyperion image taken on 30 October 2005, has been used. In this research Hyperspectral Imager (Hyperion) and Advanced Land Imager (ALI) sensor images have been used onboard EO-1 satellite. EO-1 was launched in November 2000 to demonstrate the feasibility and to evaluate the performance of a variety of innovative sensor, spacecraft, and operational technologies for future space missions (USGS, 2005). The one-year technology-demonstration mission was designed to bridge the period between 20-year-old Landsat technology and the planned (2006) Landsat Data Continuity Mission (LDCM). The spacecraft carries three instruments, of which two continue to be operated: the Advanced Land Imager (ALI), which acquires data in 10 spectral bands that cover a wavelength range from the visible to short wave infrared and resolutions similar to Landsat; and, Hyperion, which acquires data in 220 bands, 10-nanometer-wide bands from 0.4 to 2.6 micrometers (USGS, 2005). The instrument captures 256 spectra over a 7.5 km-wide swath perpendicular to the satellite motion (Kruse 2002). Figure 1 shows a HYPERION image of the test site, and its land cover. As the test site, area of interest is located in Khanbogd soum (subprovince), Umnugovi aimag (province). The area of interest covers from east to west about 6 km and from the north to south about 6.7 km.

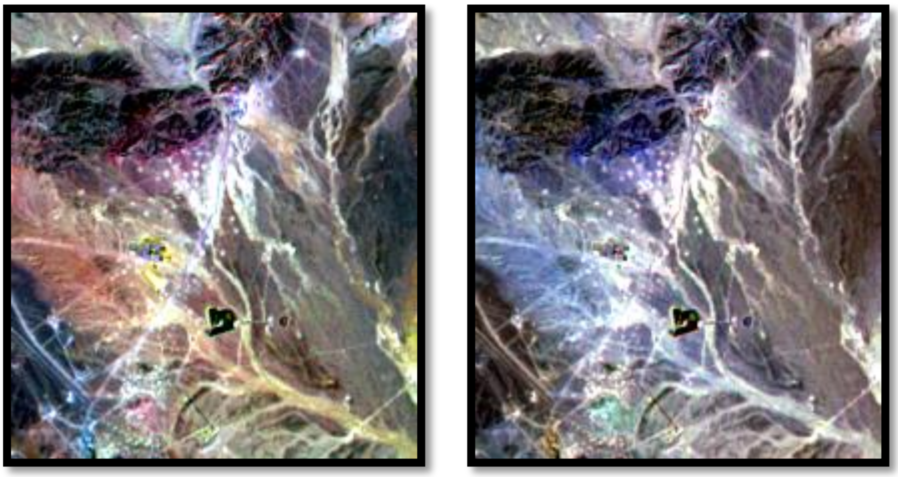

Figure 1. Hyperion image of the test area

a) RGB- 23, 53, 143 b) RGB-54, 105, 157 bands

## 3. FEATURE EXTRACTION

Firstly, Hyperion satellite images have been analyzed in terms of radiometric values, because some of the water absorption bands and some other visible, infrared, shortwave infrared bands of the image had zero values. When these bands excluded, the original satellite dataset was decreased from 242 bands to 129. For the feature extraction principal component transformation (PCT) and spectral knowledge have been used.

- Feature extraction using principal component transformation (PCT). The PCT is a statistical technique that transforms a multivariate data set of intercorrelated variables into a set of new uncorrelated linear combinations of the original variables, thus generating a new set of orthogonal axes (Richards and Xia, 1999). It is also a data compression technique used to reduce the dimensionality of the multidimensional datasets and helpful for image encoding, enhancement and multitemporal dimensionality (Pohl and Van Genderen 1998). PCT has been performed using all available bands and the result showed that the first three principal component (PC) contained 97.86% of the overall variance (83.5%, 11.56%, 3.9% for the PC1, PC2 and PC3, respectively). The visual inspection of a PC4 that contained only 0.5% of the overall variance, indicated that it contained noise. Likewise, the other PCs contained noise from the total data set. A colour image created by the use of the first three PCs is shown in Figure 2a.
- Application of spectral knowledge of the classes of interest. Nowadays, application of a knowledge-based approach has more and more usage in spectral classification of RS images. The knowledge in image classification can be represented in different forms depending on the type of knowledge and necessary of its usage. In our case, spectral knowledge of the classes of objects was used for selection of the features and it is defined on the basis of the general spectral characteristics of the classes of objects and the available spectral knowledge. Initially, the pixels representing the selected classes have been chosen from different parts of the image. Then, the statistics of these pixels was defined and plotted in a feature space and the bands which demonstrated the maximum separabilities were chosen for a further analysis (i.e., bands 24, 53, 143). A colour image created by the use of this method is shown in Figure 2b.

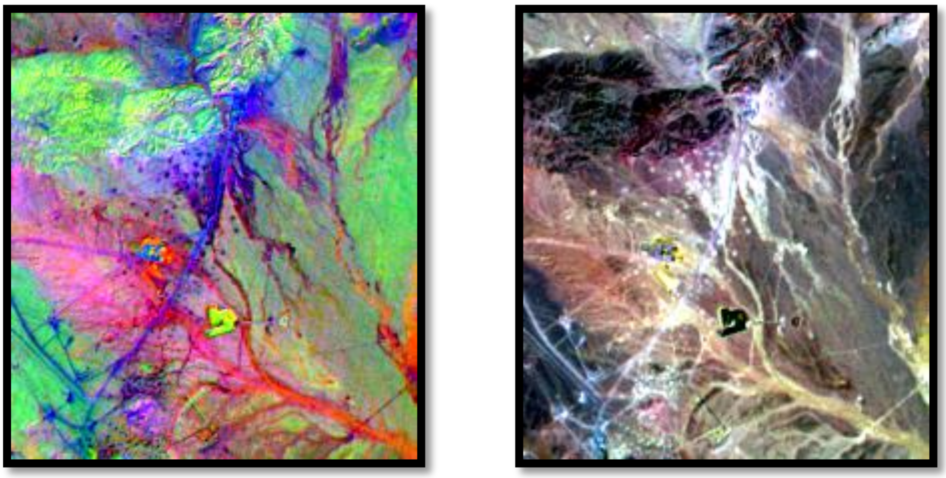

Figure 2. a) Image created by the PCA method.  b) Image created by the use of bands 24, 53, 143.

## 4. CLASSIFICATION OF THE FEATURES

Initially, to define the sites for the training signature selection, areas of interest (AOI) representing the available five classes (residential, dirt road area, soil, rock, bare soil, mine area) have been selected from the hyperspectral image. The separability of the training signatures was firstly checked in feature space and then evaluated using Jeffries–Matusita distance. The values of Jeffries–Matusita distance range from 0 to 2.0 and indicate how well the selected pairs are statistically separate. The values greater than 1.9 indicate that the pairs have good separability (Amarsaikhan *et al.* 2010). After the investigation, the samples that demonstrated the greatest separability were chosen to form the final signatures. The final signatures included about 39–545 pixels.

For the actual classification, a maximum likelihood classification and spectral angle mapper methods have been used. The maximum likelihood classification is the most widely used statistical classification technique, because a pixel classified by this method has the maximum probability of correct assignment. The spectral angle mapper is one of the most widely used hyperspectral classification techniques and it uses an *n*-dimensional angle to match pixels to reference spectra. The method determines the spectral similarity between two spectra by calculating the angle between the spectra, treating them as vectors in a space with dimensionality equal to the number of bands (Amarsaikhan *et al.* 2011). The final classified images are shown in figure 3.

As could be seen from figure 3, the spectral angle mapper classification result of the PC image give the worst results, maximum likelihood classification using PC bands shows superior result. Comparing 2 classification results obtained by the use of the spectral knowledge and PC bands (figure 3.c.d), that the performance of the maximum likelihood classification using PC bands was better than the other method.

For the accuracy assessment of the classification results, the overall performance has been used. This approach creates a confusion matrix in which reference pixels are compared with the classified pixels and as a result an accuracy report is generated indicating the percentages of the overall accuracy (Amarsaikhan *et al.* 2011). As ground truth information, different AOIs containing 1064 purest pixels have been selected. AOIs were selected on a principle that more pixels to be selected. The overall classification accuracies for the selected classes were 77.63 % and 79.04%, for the results of the maximum likelihood classification using PC bands and spectral knowledge and spectral angle mapper classification, 35.43% and 53.1% for the results of the PC bands and spectral knowledge.

## 5. CONCLUSION

The overall idea of the research was to test and compare two different approaches for feature extraction in a hyperspectral image classification. For this aim, initially, the hyperspectral HYPERION image was analyzed in terms of radiometric quality and the water absorption bands and some other bands with zero values were excluded. For the actual feature extraction, principal components analysis and spectral knowledge were used and for each case, 3 spectrally separable bands were defined. Then, these outputs were classified using a maximum likelihood classification and spectral angle mapper methods. As could be seen from the results of the classifications, the maximum likelihood classification in combination with the PC bands produced a superior result in comparison with other methods. Also, thorough analysis of the HYPERION image indicated that hyperspectral images could be used for an improved land cover mapping and differentiation of the classes having similar spectral characteristics.

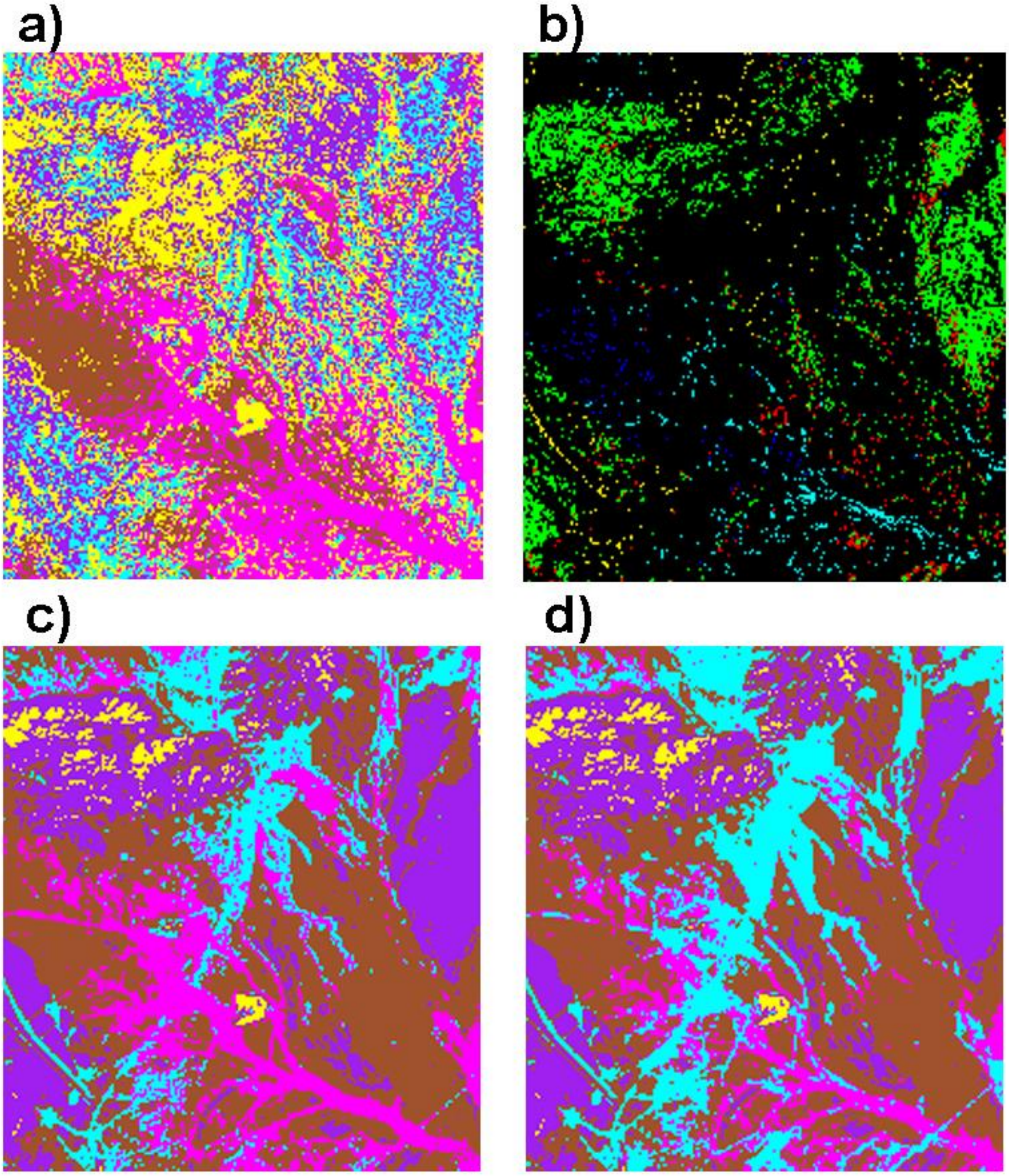

Figure 3. Comparison of the classification results (cyan-urban, dirt road, yellow-mine area, sienna-soil, magenta-bare soil, purple-bedrock). Classified images using and spectral angle mapper classification (a,b), maximum likelihood classification (c,d) defined by spectral knowledge, PC bands.

## 6. REFERENCES

Amarsaikhan, D. and Ganzorig, M., 1999. Different approaches in feature extraction for hyperspectral image classification. In: Proceedings of the 20th Asian Conference on RS, Hong Kong, pp.434-438.

Amarsaikhan, D., Blotevogel, H.H., van Genderen, J.L., Ganzorig, M., Gantuya, R. and Nergui, B., 2010. Fusing high resolution TerraSAR and Quickbird images for urban land cover study in Mongolia. International Journal of Image and Data Fusion, Vol.1, No.1, pp.83-97.

Amarsaikhan, D., Enkhmanlai, A., Bayarbaatar, D. and Enkhjargal, D., 2011. Classification of hyperspectral images of Mongolia. Informatics, No.11, pp.1-8.

Jia, X., Richards, J.A. and Ricken, D. E., 1999. Remote sensing digital image analysis: an introduction. Berlin: Springer-Verlag.

Keshava, N. and Mustard, J.F., 2002. Spectral unmixing. IEEE Signal Processing Magazine, 19, pp.44−57.

Kruse, F.A., 2002, Comparison of AVIRIS and Hyperion for hyperspectral mineral mapping, 11th JPL Airborne Geoscience Workshop, 4-8 March 2002, Pasadena, California.

Plaza, A., Benediktsson, J.A., Boardman, J.W. and Brazile, D., 2009. Recent advances in techniques for hyperspectral image processing. Remote Sensing of Environment, 113, pp.110–122.

Pohl, C., and Van Genderen, J. L., 1998. Multisensor image fusion in remote sensing: concepts, methods and applications. International Journal of Remote Sensing, 19, 823–854.

Richards, J. A., 2005. Analysis of remotely sensed data: The formative decades and the future. IEEE Transactions on Geoscience and Remote Sensing, Volume 43, pp.422−432.

Richards, J. A. and Xia, S., 1999. Remote Sensing Digital Image Analysis—An Introduction, 3rd edn (Berlin: Springer-Verlag).

Smith, M.O., Ustin, S.L., Adams, J.B. and Gillespie, A.R., 1990. Vegetation in deserts: II. Environmental influences on regional abundance. Remote Sensing of Environment, Volume 31, pp.27−52.

Yang, C., Everitt, J.H. and Johnson, H.B., 2009. Applying image transformation and classification techniques to airborne hyperspectral imagery for mapping Ashe juniper infestations. International Journal of Remote Sensing, Volume 30, Issue 11, pp.2741-2758.