

# CLASSIFICATION OF VEGETATION AND SOIL USING IMAGING SPECTROMETER DATA

J. H. Lumme

Institute of Photogrammetry and Remote Sensing, Helsinki University of Technology, P.O.Box 1200, FIN-02015 HUT, Finland  
- juho.lumme@hut.fi

Commission VII, WG VII/1

**KEY WORDS:** Remote Sensing, Hyper Spectral, Land Cover, Classification, Pushbroom

## ABSTRACT:

Monitoring the Earth using imaging spectrometers has necessitated more accurate analyses and new applications to remote sensing. New algorithms have been developed for hyperspectral data classification lately, but also traditional classification algorithms have often been used. This study compares different classification algorithms for classification of vegetation using imaging spectrometer data. The test area located in southern Finland was imaged by an AISA airborne imaging spectrometer using 17 visual and near infrared bands. The area included lakes, rural areas, cultivated fields and forests. The area was classified into seven different vegetation and soil types. The effects of various classification algorithms and different training areas were investigated. Besides, the reflectance spectra of different plants were examined and compared under varying illumination. Spectral Angle Mapper (SAM), Spectral Correlation Mapper (SCM) and Spectral Unmixing algorithms developed for hyperspectral data were used in the classification. Besides, the data was classified using conventional algorithms as Minimum Distance and Maximum Likelihood classifiers that have often been used for multispectral data classification in the past. The dimension of the data was decreased by principal component analysis before using conventional classifiers. Reference spectra for SAM, SCM and Spectral Unmixing were collected from the training sites of the data. Two methods were used in gathering the reference spectra. The reference spectra were chosen from the reflectance of individual image pixels or they were calculated from pixels of the training sites. If individual pixel was chosen accurately, it led to better classification results. Maximum Likelihood classifier led to good results as well, but it requires more computation time. The overall accuracy of the Maximum Likelihood classification was 91 percent, but the results deteriorated under varying illumination. SAM and SCM were faster and they led to better classification results in poor illumination. The hardest part in Spectral Unmixing classification was finding suitable reference spectra from mixed pixels. When the essential spectra were found, the classification led to good results, although the results varied between different classes.

## 1. INTRODUCTION

Imaging spectrometers have been developed very rapidly over past decades. They have more channels with better spatial and spectral resolution. Individual bands are only a few nm wide while the spatial resolution is good as well. Besides, computer number-crunching power, data-transfer rate and storage capacity have increased considerably in recent years. This has made it possible to handle and analyse larger data sets acquired by an imaging spectrometer. There are some new algorithms being developed to hyperspectral data classification at the moment.

Most hands-on applications of imaging spectrometry relate to a sensor mounted on aeroplane. Satellite sensors do not have as good spectral and spatial resolution as aeroplane sensors, which may have hundreds of channels with one meter spatial resolution. Nowadays, satellite sensors tend to have tens of channels with spatial resolution of tens or hundreds meters.

Traditional algorithms like Maximum Likelihood or Minimum Distance have also been used in hyperspectral data classification. Unfortunately, they tend not to work properly with hyperdimensional data. Hyperspectral image data consists of hundreds of channels and it leads to extended run time. When channel number doubles, run time squares. Highly correlated channels may even crash the classification programme. For these reasons effective feature extraction

methods are needed to reduce hyperspectral feature space dimension.

Feature extraction methods may remove small differences in the spectra between various materials. Sometimes these differences characterize different materials and play an important role in the classification. When using a hyperspectral classification algorithm like Spectral Angle Mapper the feature extraction is not always essential. This way the whole reflectance spectrum of the material is taken into account. For instance, the effect of different soil types on the reflectance spectrum of vegetation is slight and it comes up only with some specified wavelength bands.

Different institutions use imaging spectrometer data for various research projects in Finland. For example, workers in the Finnish Forest Research Institute are researching imaging spectrometer data for inventory of forest resource (Mäkisara and Tomppo, 1996). Hyperspectral data is also used for monitoring the quality of water (Kallio et al., 2001) and atmosphere at the Finnish Environment Institute.

The goals of this study were to investigate the suitability of AISA imaging spectrometer data for vegetation and soil classification. Different classification algorithms were compared using various reference data and illumination. Besides, reference spectra of different materials were analysed and material identification by its spectrum was investigated.

There were some different methods in gathering the reference spectra and these methods were tested and their results were compared. First, the reference spectrum was chosen from the reflectance of individual and exactly specified image pixels. Next, the reference spectrum was calculated from pixels of the training sites. Average, median and mode methods were used to calculate the reference spectrum.

## 2. METHODS

### 2.1 Classification algorithms

**2.1.1 Spectral Angle Mapper (SAM):** Reflectance spectrum of individual pixel may be discussed as an  $n$ -dimensional vector, where  $n$  is the number of image channels. Each vector has certain length and direction. The length of the vector represents brightness of the target while the direction represent spectral feature of the target. Variation in illumination mainly effects changes in the length of the vector. Therefore, classification is based on the direction of the vector. (Kruse et al., 1993)

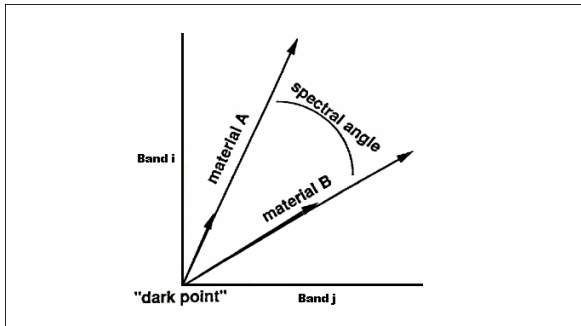


Figure 1. The spectral angle between material A and B in two channels case (Kruce et al., 1993).

Classification is done by comparing the spectral angles (Figure 1) between the reflectance spectrum of the classified pixel and the reference spectra obtained from training data or spectral libraries. Each pixel will be assigned to the class according to the lowest spectral angle value.

**2.1.2 Spectral Correlation Mapper (SCM):** SAM cannot distinguish between negative and positive correlations because only the absolute value is considered. SCM is generated as an improvement on the SAM. SCM algorithm is very similar to SAM. The difference between the algorithms is that SCM standardizes the vectors of the reflectance spectrum before it calculates the spectral angles. (Carvalho and Meneses, 2000)

**2.1.3 Spectral Unmixing:** An area assigned by a single pixel of remote sensing image usually contains a lot of different materials. These materials are mixed together and the pixel reflectance observed by sensors is a combination of reflectance of individual materials. To get more information from a single pixel the proportions of these materials can be approximated using a spectral mixing model (Boardman, 1994). Using Spectral Unmixing model the mixed pixel can be reconstructed from known spectra in the image or the mixed pixel can be divided into components.

### 2.2 Test site and field measurements

Aerial measurements with an AISA airborne imaging spectrometer were made at Lammi in southern Finland. The area contains mainly lakes, rural areas, cultivated fields and coniferous and deciduous forests. Size of the test area was about 50 kilometres long and 2 kilometres wide.

Geological Survey of Finland had done fieldworks in 1999 and 2000. Training area inventory included approximately 250 training areas. For example, the primary sources of reflection, the vegetation and the soil class were specified in field inventory.

### 2.3 AISA hyperspectral data

AISA data was acquired from an aeroplane on September 1999. Six strips were flown and raw data was gathered to the actual hyperspectrum image and geometric correction was performed. Finnish Forest Research Institute preprocessed the data.

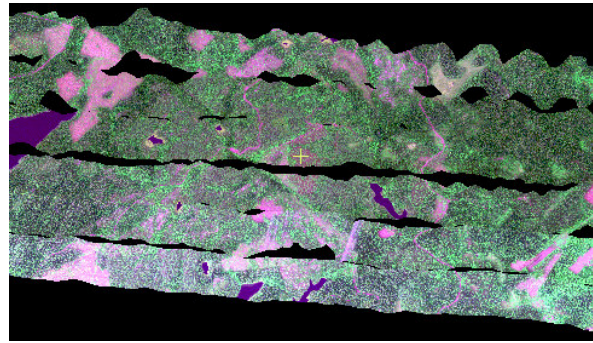


Figure 2. A part of the AISA image (Ruohomäki at al., 2002).

The pixel size of the image was 1.1 meter. Image contained 17 visible and near infrared channels. The weather took a turn for the worse during the flight and the illumination was low in last strips (Figure 2).

### 2.4 Gathering the reference spectra

Classification algorithms like SAM use reflectance spectra as reference data of classes. Reference spectra are measured from pure and single image pixels or larger training areas. The quality of reference spectra is an important factor that defines the classification results. Therefore, different methods were used in gathering the reference spectra. Besides reflectance spectra of different plants were examined and compared under varying illumination to describe how different classes may be distinguished.

Well-known pure pixels of hyperspectral image were used directly as reference spectra. In that case, the reflectance of the pixel must be derived only from material that is defined by the pixel. If there is variation in reflectance features inside one class, more reflectance spectra will be needed.

Several methods were used to calculate reference spectra from the training areas in this study. At first, average values of pixel values were calculated for every 17 channels. By choosing median of pixel values was the second method to determine

reference spectra. The third method used mode values as a reference. It means most frequently occurring value.

Smaller test area (Figure 3) with 100 pixels was chosen for a more detailed analysis. Reference spectra were derived from this area (Figure 4). The area represents a dense spruce forest and it includes spruce trees, shadows and bedrock of the forest.

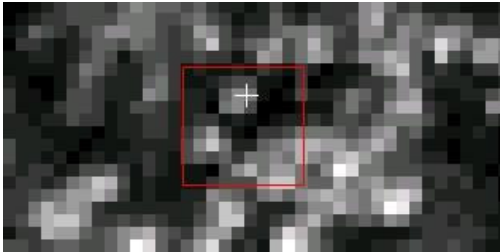


Figure 3: 10 x 10 pixel test area of hyperspectral image. Bright pixels are sunny spruce crowns. Dark pixels are shadows between spruces.

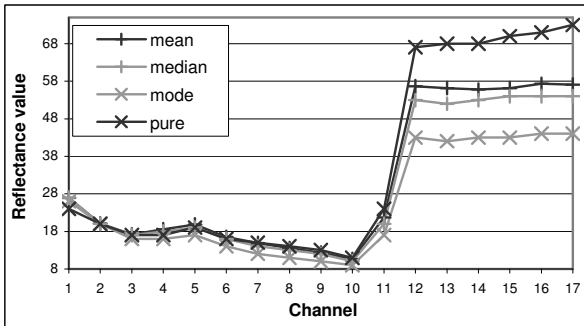


Figure 4: Reference spectra calculated by mean, median and mode method and pure spruce spectrum.

Besides, bright and pure pixels were chosen to represent the reflectance spectra of spruce. Spectra were compared with each other and the most descriptive spectrum was selected. Values were not calculated from larger group of pixels because it tends to generalize or flatten the shape of the spectrum. Most descriptive spectrum was considered as pure spruce spectrum. Pure spectrum and spectra calculated by different methods were used as reference spectra and spectral angles (Figure 1) between the each pixel of the area (Figure 3) and reference spectra (Figure 4) were calculated.

### 3. RESULTS AND DISCUSSION

#### 3.1 Reference spectra analysis

Comparison of different methods in gathering the reference spectra were analysed using the spectral angles. Spectral angle images (Figure 5) were stored. Image colours were inverted to help the interpretation of the spectral angle images. Therefore bright image pixels represent small values of spectral angle and dark pixels represent wide spectral angles. Small spectral angles mean that the reference spectrum describes the study area well.

The more there are bright pixels in the spectral angle images the better choice that reference spectrum is for describing the test area. These results are significant when choosing reference spectra for example Spectral Angle Mapper classification.

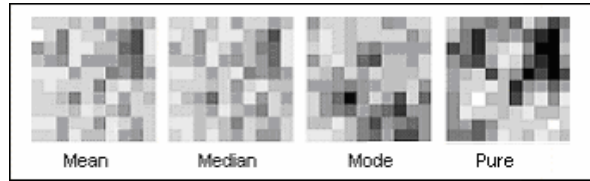


Figure 5: The spectral angles between the small test area pixels and the reference spectra.

Figure 5 shows that calculating reference spectrum by using mean values gives slightly better results than median. By comparing the test area image and the spectral angle images we can see that reference spectra of average and median method cannot define dark pixels of the test area. The smallest values of the spectral angles were found from the image pixels between sunny crowns and shadows.

The spectral angle values of the reference spectrum calculated with the mode method were high and especially mode method gave bad results in the case of sunny crown pixels. Pure pixel method defined reference spectrum that was outstandingly good in case of bright crown pixels. Anyway, the results were poor in the dark pixel cases as expected.

#### 3.2 Classification

Smaller area was selected from the AISA data and classified using several algorithms. The area was classified into the following different vegetation and soil types: roads, buildings, cornfield, threshed cornfield, sugar beet, deciduous and coniferous forests. Results of different classification algorithms were compared. Besides, effects of different training areas and variation in illumination were investigated.

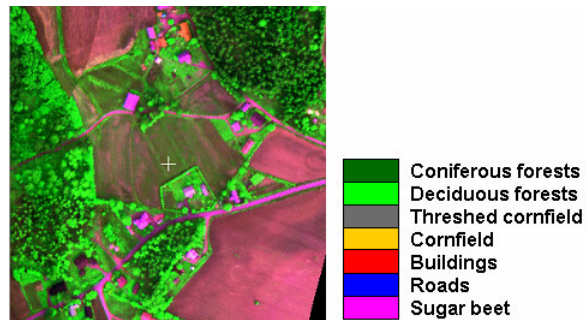


Figure 6: The image of the test area that was used in the classifications and the colour codes for the classification results.

Spatial resolution of the data was good and there was some variation in the values of the training site pixels. Therefore several reference spectra per each training site were used in the Spectral Angle Mapper and Spectral Correlation Mapper classifications. The reference spectra were gathered from the same training sites used in the Minimum Distance and

Maximum Likelihood classifications. Principal component analysis was used to reduce the dimension of the data. Four principal components were used in the Minimum Distance and Maximum Likelihood classification. Other classification methods were carried out fast but Maximum Likelihood algorithm needed more computation time.

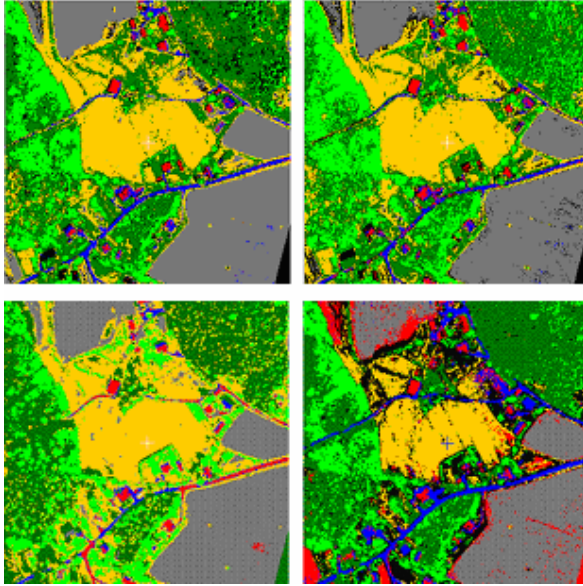


Figure 7: Classification results of Spectral Angle Mapper (upper left), Spectral Correlation Mapper (upper right), Minimum Distance (lower left) and Maximum Likelihood (lower right).

Average and overall accuracies were calculated using test areas (Figure 8). The average accuracy is the average of the accuracies for each class. Overall accuracy is a similar average with the accuracy weighted by the proportion of test samples. The Maximum Likelihood algorithm led to the best results. Spectral Angle Mapper and Spectral Correlation Mapper led to good results as well but the results of Minimum Distance were worse.

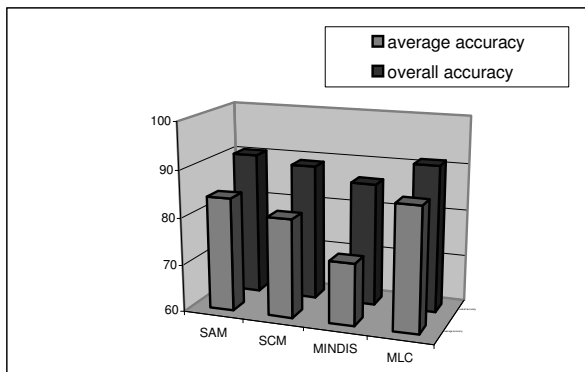


Figure 8: The average and overall accuracies of the Spectral Angle Mapper, Spectral Correlation Mapper, Minimum Distance and Maximum Likelihood classification.

**3.2.1 Variation in illumination:** Variation in illumination affected more strongly when the Maximum Likelihood and Minimum Distance were used. The results of the classification deteriorated fast while the Spectral Angle Mapper and the Spectral Correlation Mapper were better in that case.



Figure 9: The image of the test area with variance in illumination (up) and the classification results of Spectral Angle Mapper (middle) and Maximum Likelihood (down).

**3.2.2 Spectral Unmixing:** The original pixel size of the hyperspectral image was enlarged ten times its original size because a square meter size pixel is usually represented by only one manmade or vegetation class. Finding suitable reference spectra was the hardest part in Spectral Unmixing classification. The proportions of different materials were approximated from each image pixel and algorithm generated images where pixel values represented the proportions of different materials (Figure 10). Proportions were calculated from 0 to 100 per cent and bright pixels meant larger proportion.

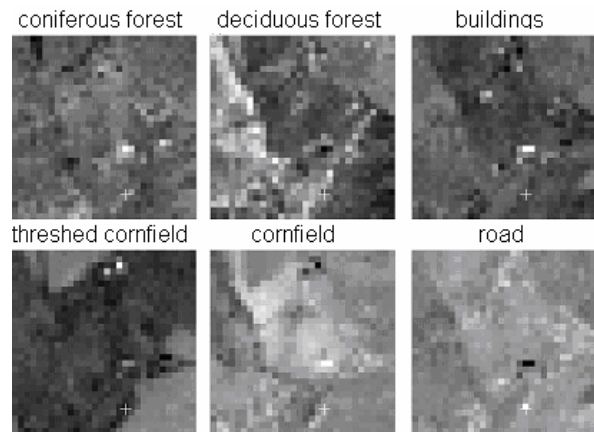


Figure 10: The proportion images of the Spectral Unmixing Classification.

The proportion images were compared to the original image (Figure 6) and there was good association between the original

image and the proportion images of threshed cornfield, deciduous forests, cornfield and buildings. For example, the brighter the pixel of the buildings proportion image is the larger the proportion of the class is in the pixel. It means that large buildings have brighter pixel. The results for coniferous forest and road proportion images were worse.

#### 4. CONCLUSIONS

Generally speaking the classification results were good. For example, Maximum Likelihood classifier led to good results (overall accuracy was about 91 percent), but it requires more computation time. Spectral Angle Mapper and Spectral Correlation Mapper were faster and they led to better classification results in poor illumination. The results of Minimum Distance classification were poor. Spectral Unmixin algorithm worked in some cases. It produced good proportion images for threshed cornfield, deciduous forests, cornfield and buildings but coniferous forest and road did not work. The suitable reference spectra for the mixed pixel were hard to find. There was too much variation in the pixel values of the same class. This study shows clearly that it is worthwhile to pay attention to different methods when the reference spectra are calculated.

#### 5. REFERENCES

- Boardman, J., 1994. Geometric mixture analysis of imaging spectrometry data. In: *International Geoscience and Remote Sensing Symposium, IGARSS '94, Surface and atmospheric remote sensing: technologies, data analysis and interpretation*. California, USA, Vol. IV, pp. 2369-2371.
- Carvalho, O. & Meneses, P., 2000. Spectral Correlation Mapper (SCM): An Improving Spectral Angle Mapper. In: *Ninth JPL Airborne Earth Science Workshop*. Pasadena, CA, JPL Publication 00-18, pp. 65-74.
- Kallio, K., Kutser, T., Hannonen, T., Koponen, S., Pulliainen, J., Vepsäläinen, J., Pyhälähti, T. 2001. Retrieval of water quality from airborne imaging spectrometry of various lake types in different seasons. *Science of the Total Environment*, 268, pp. 59-77.
- Kruse, F., Lefkoff, A., Boardman, J., Heidebrecht, K., Shapiro, A., Barloon, P. & Goetz, A. 1993. The spectral image processing system (SIPS) - interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44(2-3), pp. 145-163.
- Mäkisara, K. & Tomppo E. 1996. Airborne imaging spectrometer in National Forest Inventory. In: *International Geoscience and Remote Sensing Symposium, IGARSS '96, Remote sensing for a sustainable future*. New York, USA, Vol. II, pp. 1010-1013.
- Ruohomäki, T., Törmä, M., Rainio, H., Lumme, J. & Mäkilä, M. 2002. Soil types, vegetation and AISA-imagery in Lammi test area. *Geological Survey of Finland, GTK/RS/2002/6*