Fourth Annual JPL Airborne Geoscience Workshop, Arlington, Virginia, October 25-29, 1993.

Classification of High Dimensional Multispectral Image Data*

Joseph P. Hoffbeck and David A. Landgrebe

School of Electrical Engineering
Purdue University
West Lafayette, IN 47907

1. Introduction

A method for classifying high dimensional remote sensing data is described in this paper. The technique uses a radiometric adjustment to allow a human operator to identify and label training pixels by visually comparing the remotely sensed spectra to laboratory reflectance spectra. Training pixels for materials without obvious spectral features are identified by traditional means. Features which are effective for discriminating between the classes are then derived from the original radiance data and used to classify the scene. This technique is applied to Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) data taken over Cuprite, Nevada in 1992, and the results are compared to an existing geologic map. This technique performed well even with noisy data and the fact that some of the materials in the scene lack absorption features. No adjustment for the atmosphere or other scene variables was made to the data classified. While the experimental results compare favorably with an existing geologic map, the primary purpose of this research was to demonstrate the classification method, as compared to the geology of the Cuprite scene.

2. The Scene

The data set used to demonstrate the technique is from the 1992 AVIRIS flight over the Cuprite Mining District in southwestern Nevada. The raw data consist of 220 spectral bands covering the range 0.4 to 2.5 μm. The site, which lies on the east side of U.S. Highway 95, has very little vegetation and several exposed minerals including alunite, buddingtonite, kaolinite, and quartz.

3. Identifying Training Pixels using Absorption Features

The reflectance spectra of four minerals as measured in the laboratory are shown in Figure 1 (Goetz et al., 1985). While the absorption features can be easily seen in the reflectance spectra measured in the laboratory, they can be quite difficult to see in the remotely sensed radiance spectra; therefore, the log residue method (Green et al., 1985) was used to adjust the shape of the radiance spectra to be more similar to the laboratory reflectance spectra. The log residue spectra were visually compared to the laboratory spectra, and 730 pixels of alunite, 71 pixels of buddingtonite, 232 pixels of kaolinite, and 385 pixels of quartz were identified in the Cuprite scene. Figure 2 shows the log residue spectrum of the average of each class.

Some pixels in the scene were mixtures of alunite and kaolinite. Since these pixels were not selected as training pixels, the final classification represented the dominant mineral in each pixel. If one were interested in mapping mixed pixels, training pixels for mixed alunite and kaolinite could be used to define an additional class.

* Work leading to this paper was funded in part by NASA Grant NAGW-925.

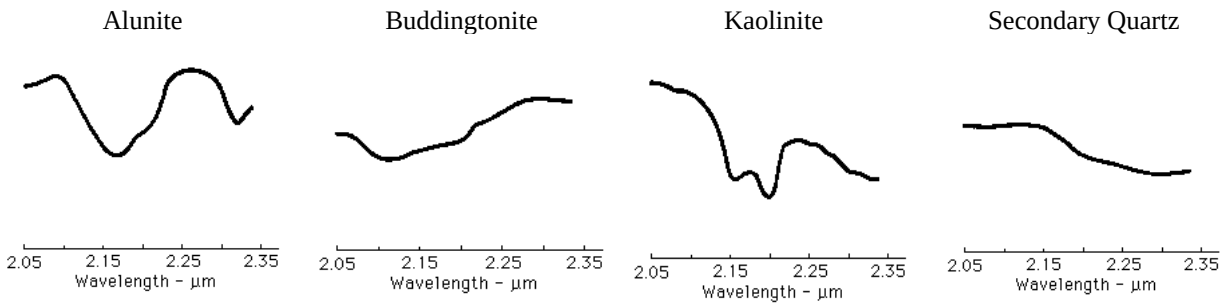

Figure 1. Laboratory Reflectance Spectra of four minerals (Goetz et al., 1985).

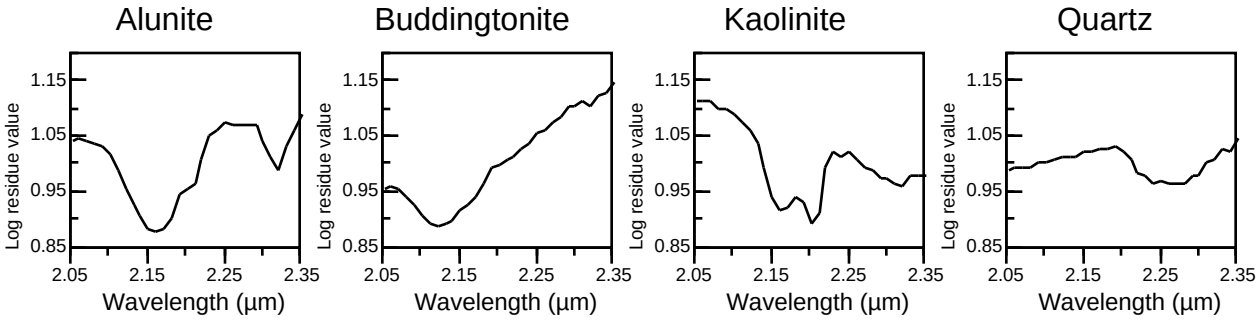

Figure 2. Log Residue of the Average of Each Class.

4. Initial Classification to Identify Other Classes

For materials that lack strong absorption features or whose reflectance spectra is unknown, it may be impossible to directly identify training pixels by comparison to laboratory reflectance spectra. Training pixels for these classes were identified as follows. Discriminant analysis (Fukunaga, 1990) was run on the raw radiance data using the four classes with known absorption features and using all 220 AVIRIS bands (0.4 to 2.5 µm). A maximum likelihood classifier was used to classify the first eight features and a probability map was produced. Areas of low likelihood suggested the location of further classes. Training pixels for argillized, tuff, and alluvium were then located in these areas.

Note that since classes like argillized, tuff, and alluvium do not have known absorption features, additional information was required to determine their identity. In this case the names of the classes argillized, tuff, and alluvium were determined by comparison with the geologic map in Figure 4 (Abrams et al., 1977). Had this information not been available, the training samples still could have been located and the classes mapped, but the names of the classes would not be known.

5. Feature Extraction and Classification

Discriminant analysis was run using all 8 classes and all 220 bands of the raw radiance data. Using the first seven discriminant features, the thematic map in Figure 3 was produced with a maximum likelihood classifier. A threshold was used classify pixels that were very unlikely to belong to any of the training classes as background. The classes background, alluvium, and tuff were all displayed as white in Figure 3.

The result can be compared to the geologic map in Figure 4* (Abrams et al., 1977). This map divides the Cuprite region into a silicified zone which contains abundant quartz, an opalized zone which contains opal, alunite, and kaolinite, and an argillized zone. Comparing Figure 3 to the geologic map shown in Figure 4, note first that the region classified as quartz corresponds to the silicified zone in the geologic map. Note in addition to the central quartz zone, there two small quartz zones to the upper left and the lower left of the central zone that are similar in

* No attempt was made to adjust the geometry of the AVIRIS data to any geographic coordinate system, thus the geometry of Figures 8, 9, and 10 may not coincide.

both maps. Next note that the regions classified as alunite and kaolinite lie in the opalized zone in the geologic map. The region classified as argillized corresponds to the region in the geologic map with the same name. Buddingtonite is not shown in the geologic map. Other researchers have classified this scene (Kierein-Young et al., 1989, Hook et al., 1990, Kruse et al., 1990).

The classification accuracy is difficult to assess quantitatively without knowledge of the actual dominant mineralogy for each pixel in the scene. Two indirect methods to estimate the accuracy of a maximum likelihood classification are the resubstitution method which is optimistically biased, and the leave-one-out method which provides a lower bound on the accuracy (Fukunaga, 1990). If the training pixels were representative of the classes in the scene, the true accuracy for the pixel classification would lie somewhere between these two values. Both the resubstitution accuracy and the leave-one-out accuracy for the classification in Figure 3 were 99.6%.

The fact that this method of classifying high dimensional remote sensing data is relatively insensitive to noise is demonstrated by noting that the final classification changed very little regardless of whether or not the water absorption bands were included in the analysis. The reflected radiation in these bands (1.36 - 1.41 µm and 1.82 - 1.93 µm) is completely absorbed by the atmosphere, and so the data in these bands contains only noise. When the water absorption bands were included the resubstitution accuracy and the leave-one-out accuracy were both 99.6% as mentioned above. When the water absorption bands were not included the resubstitution accuracy was 99.7% and the leave-one-out accuracy was 99.6%.

6. Conclusion

In this paper, a method has been demonstrated for analyzing a data set of high spectral dimensionality. Such high dimensional data not only makes possible the use of narrow spectroscopic features where they are known to exist, but is also able to make available the inherently higher information content of such data as predicted by signal theory principles. Training pixels for materials with strong absorption features were located in the data using the log residue method to adjust the radiance spectra to resemble the reflectance spectra. Training pixels for materials without strong absorption features or known reflectance spectra were located using other knowledge such as photo interpretation, ground observations, etc. Features maximally effective in discriminating between the classes so defined were then computed from the raw radiance data using the discriminant analysis method, and these features were classified using the maximum likelihood classifier.

This method generated good results even with noisy data and with classes that lacked strong absorption features. This method effectively combined the human operator's knowledge of chemical spectroscopy with the power and robustness of the statistical classifier to perform the classification, greatly reducing the dependence of the analysis process on both reflectance and wavelength calibration and on high signal-to-noise-ratio. It also does not require ground truth in the sense of observations taken from the ground, and thus can be applied to the analysis of data gathered from remote areas.

Bibliography

Abrams, M.J., R.P. Ashley, L.C. Rowan, A.F.H. Goetz, and A.B. Kahle, 1977, "Mapping of Hydrothermal Alteration in the Cuprite Mining District, Nevada, Using Aircraft Scanner Images for the Spectral Region 0.46 to 2.36 µm," *Geology*, vol. 5, no. 12, pp. 713-718.
Fukunaga, K., 1990, *Introduction to Statistical Pattern Recognition*, Second Edition, Academic Press, New York, pp. 220-228, 445-455.

Goetz, A.F.H. and V. Srivastava, 1985, “Mineralogical Mapping in the Cuprite Mining District, Nevada,” *Proceedings of the Airborne Imaging Spectrometer Workshop*, JPL Publication 85-41, pp. 22-31.

Green, A.A. and M.D. Craig, 1985, "Analysis of Aircraft Spectrometer Data With Logarithmic Residuals." *Proceedings of the Airborne Imaging Spectrometer Workshop*, JPL Publication 85-41, pp. 111-119.

Hook, S.J. and M. Rast, 1990, "Mineralogic Mapping Using Airborne Visible Infrared Imaging Spectrometer (AVIRIS) Shortwave Infrared (SWIR) Data Acquired Over Cuprite, Nevada", *Proceedings of the Second AVIRIS Workshop*, JPL Publication 90-54, pp. 199-207.

Kierein-Young, K.S. and F.A. Kruse, 1989, "Comparison of Landsat Thematic Mapper and Geophysical and Environmental Research Imaging Spectrometer Data for the Cuprite Mining District, Esmeralda and Nye Counties, Nevada", *Digest - International Geoscience and Remote Sensing Symposium (IGARSS)* , vol. 2, pp 961-964.

Kruse, F.A., K.S. Kierein-Young, and J.W. Boardman, 1990, "Mineral Mapping at Cuprite, Nevada With A 63-Channel Imaging Spectrometer," *Photogrammetric Engineering and Remote Sensing*, vol. 56, no. 1, pp 83-92.

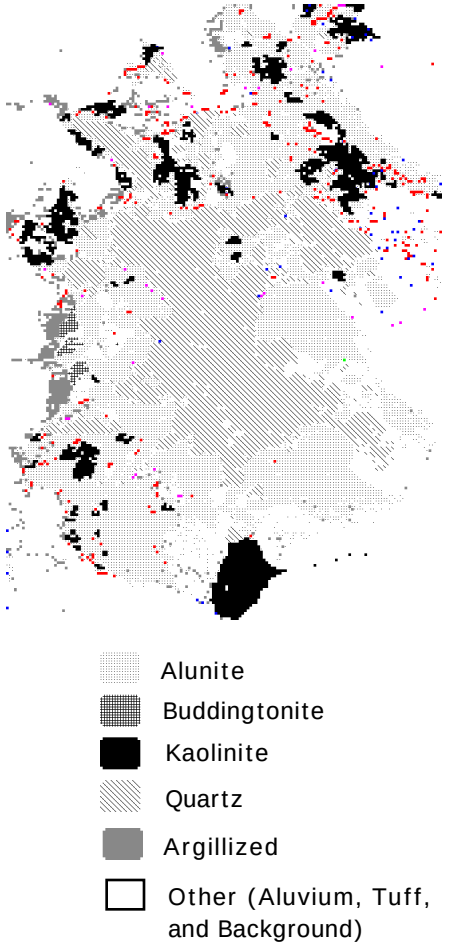

Figure 3. Classification of AVIRIS data.

Figure 4. Geological map redrawn from Abrams et al., 1977.