AUTOMATIC ENDMEMBER EXTRACTION FROM HYPERSPECTRAL DATA FOR MINERAL EXPLORATION*

R.A. Neville and K. Staenz
Canada Centre for Remote Sensing
Ottawa, Ontario, Canada

T. Szeredi and J. Lefebvre
MacDonald Dettwiler and Associates
Richmond, British Columbia, Canada

P. Hauff
Spectral International Inc.
Arvada, Colorado, USA

ABSTRACT

Linear unmixing is a commonly used technique for the analysis of imaging spectrometer data. This procedure requires the use of 'endmember' spectra which represent pure target species. One of the methods under investigation at Canada Centre for Remote Sensing for the automatic extraction of endmembers from the image cube, is called the 'Iterative Error Analysis', (IEA) method. This procedure is described, and the results of its application to a data set acquired over Nevada are presented.

1.0 INTRODUCTION

Imaging spectrometers acquire images in many, up to several hundred, contiguous spectral bands simultaneously. Such data are well suited to the identification of target materials having spectral reflectances distinguished primarily by narrow spectral features, and are particularly important in discriminating between materials for which these features are limited to the same small region of the spectrum. One application requiring such data is the mapping of clay minerals. The reflectance spectra of these minerals have narrow (down to 10 nm in width), absorption features in the 2050 nm to 2350 nm part of the spectrum. In addition, the exact wavelength location and shape of a particular absorption feature may vary for an individual mineral type depending on the chemical composition of the sample in question. Such information can be extracted from imaging spectrometer data by an expert analyst.

There is an increasing number of airborne imaging spectrometers in operation. As examples we mention the following: the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), (Vane *et al*., 1993) covering the spectral range between 400 nm and 2500 nm with 224 bands; Probe-1 with 128 bands over the range from 440 nm to 2543 nm; the Compact Airborne Spectrographic Imager (*casi*), (Anger *et al.*, 1996) with up to 288 bands in the visible and near infrared part of the spectrum, (400 nm to 1000 nm); and the Short wave infrared Full Spectrum Imager (SFSI), (Neville *et al.*, 1995) with 235 bands covering the spectral range from 1200 nm to 2450 nm.

One of the methods commonly used (Adams, *et al.*, 1986; Boardman, 1995; Szeredi, *et al.*, 1999) for the analysis of imaging spectrometer data is spectral unmixing. In general, a pixel observed by the remote sensing instrument consists of mixed materials. The hypothesis underlying *linear* unmixing is that the spectral radiance measured by the sensor consists of the radiances reflected by all of these

* Presented at the Fourth International Airborne Remote Sensing Conference and Exhibition / 21st Canadian Symposium on Remote Sensing, Ottawa, Ontario, Canada, 21-24 June 1999.

the spectral radiance measured by the sensor consists of the radiances reflected by all of these materials, summed in proportion to the subpixel area covered by each material. To the degree that this hypothesis is valid, and that the 'endmembers' are given by the reference spectra of each of the individual pure materials, and under the condition that these spectra are linearly independent, then in theory one can deduce the makeup of the target pixel by calculating the particular combination of the endmember spectra required to synthesize the target pixel spectrum. In practice there are a number of factors which act to confuse the issue. Some of these are sensor artifacts, sensor noise, atmospheric effects, solar incidence and terrain slope angles, surface roughness and other radiometric influences.

One possible source for the endmember spectra are libraries of spectral reflectances. The risk in using such library spectra in the unmixing operation is that the library spectra are rarely, if ever, acquired under the same conditions as the airborne data. The size of the particles constituting the mineral sample and the illumination conditions are but two of the variables that can have a significant effect on the resulting spectra. A better match will be obtained if the endmember spectra are taken from the image cube under analysis.

Techniques have been developed to extract endmember spectra from the remotely sensed data set, (Boardman, 1995). These in general involve the representation of the set of all scene pixels as a scatter plot in spectral space or some subspace thereof. Each pixel's spectrum, consisting of M values, corresponding to M spectral bands, can be represented as a point in an M-dimensional space. Because M is generally very large for imaging spectrometer data, a transformation, such as a Principal Component (PC) transformation is applied to the data to permit its representation in a space of lower dimensionality. Each scene spectrum maps to a point in this PC space and the set of all scene spectra constitutes a scatter plot. The endmember spectra occur at the extremities of this scatter plot. The methods most often used to acquire the endmember spectra from the data rely upon computer visualization tools which allow manipulation of this scatter plot, permitting one to project the plot onto an arbitrary plane in the PC transformed space. This permits an operator to locate the extremities, the endmember spectra, in the scatter plot. The spectra of all the pixels selected from the extremities in this scatter plot then constitute the set of endmember spectra.

These techniques are best suited to the processing of relatively small data sets, for example, of less than 10^6 pixels. Current imaging spectrometers, however, are quite capable of collecting up to 20,000 spectra per second, i.e. 10^7 Bytes per second, or 36 GBytes per hour. On even a moderately good weather day, an airborne imaging spectrometer can be expected to collect up to 150 Gbytes of data. Given these instrument capabilities, and the growing interest in the use of imaging spectrometer data for applications such as mineral exploration, it will become increasingly important to have an automatic method for extracting the endmember spectra from the data.

Several automatic techniques have been investigated at Canada Centre for Remote Sensing, (Szeredi *et al.*,1999). Of these, the most robust method is the 'Iterative Error Analysis' (IEA) method. It has been tested on image data from a number of sensors for a number of different target types. One of these data sets was acquired by SFSI over a geological remote sensing test area at Cuprite, Nevada. The IEA technique is described in Section 2.1 . The results of its application to the SFSI Cuprite data set are presented in Section 2.2, and analysed in Section 2.3 . The performance of the IEA method is summarized in Section 3.

2.0 ITERATIVE ERROR ANALYSIS

2.1 METHOD

IEA is a method which relies on the existence of relatively pure pixels. In this method, a series of linear, constrained unmixings is performed (Neville, *et al.*, 1998), each time choosing as endmembers those pixels which minimize the remaining error in the unmixed image. The algorithm stops when a termination condition is reached: typically either a certain predetermined number of endmembers, N say, is obtained, or a predetermined error tolerance is reached. The algorithm is executed directly on the spectral data. No transformation into Principal Component (PC), or other transform, space are required.

An initial vector is required to start the procedure, and for this purpose the mean spectrum, $_{mean}$, of the data is calculated. A constrained unmixing in terms of this vector is performed and the 'error image', the image of the errors remaining after the unmixing, is found. The vectors farthest from the mean spectrum(in terms of Euclidean distance) show up as the vectors with the largest error. Presumably, these vectors lie at the tip of one of the vertices of the data cloud and hence are relatively pure pixels. These pixels are averaged to form the first endmember. The averaging is performed for the purpose of decreasing the effects of outliers and noise. Another constrained unmixing is performed using this first endmember and the vectors farthest from the first endmember show up as the vectors with the largest error. Again, presumably these vectors lie at the tip of one of the vertices of the data cloud. They are averaged to form the second endmember. This process is repeated until either the unmixing error is below some threshold or a predetermined number of endmembers is reached.

In the current configuration, the operator selects the desired number of endmembers N, a number R, and an angle θ . R is the number of pixels with the largest errors in the error image after each unmixing. The vector $_{max}$, which is the spectral vector corresponding to the pixel with the single largest error, is found. That subset of the set of R vectors consisting of all those vectors which fall within an angle θ of the maximum error vector $_{max}$ is found and these are averaged to give the endmember vector. By increasing the value of R, one can reduce the effect of noise. By reducing the magnitude of the angle θ, one can improve the spectral purity of the resulting endmembers.

It is important to use *constrained* unmixing. The use of unconstrained unmixing with real data typically results in large positive and negative fractions for some pixels. Although the endmember fractions may still sum to close to 1, they are typically not positive. This results in spurious endmembers being chosen by the algorithm.

This technique has proven to be robust. In fact, even if the parameters R and θ are changed, within rather liberal limits, essentially the same endmembers are obtained. The only caution one must exercise is to avoid selecting values for R or θ which are so small that single pixel endmembers are returned; the degree to which this proviso is important is inversely proportional to the signal-to-noise ratio (SNR) of the image data. An additional advantage of this method is that the endmembers are automatically prioritized; the order in which the endmembers appear is directly related to the degree to which they are required to minimize the unmixing errors.

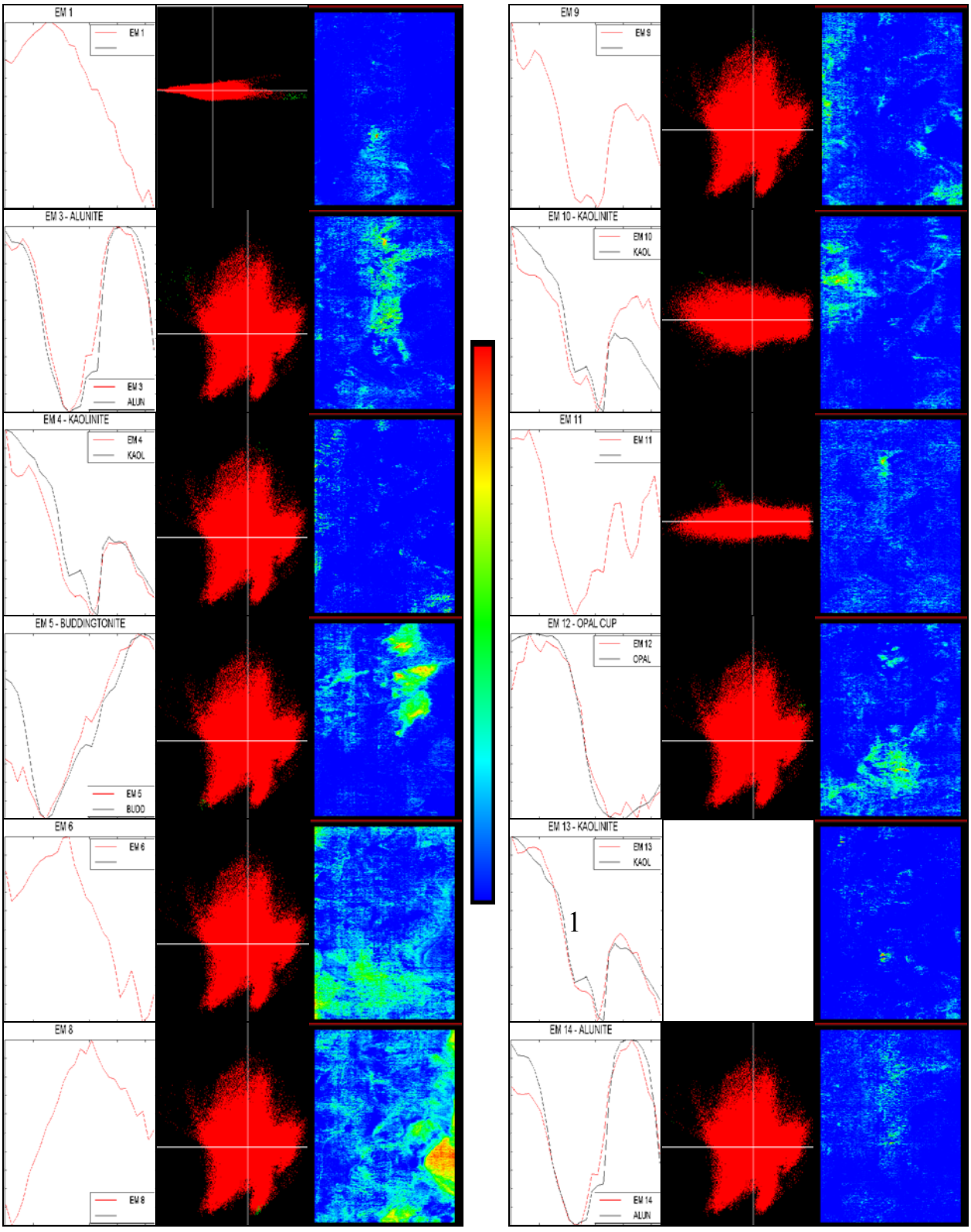
2.2 RESULTS

To demonstrate this method, a SFSI radiance cube in the Cuprite region of Nevada was processed. The radiance data was converted to surface reflectance based on a look-up table approach involving MODTRAN3, as described by Staenz and Williams, (1997). Thirty bands between 2050 nm and 2350 nm were used in the endmember extraction and unmixing analysis. Sixteen endmembers were requested from the IEA module, with the parameters $R = 50$ [pixels] and $\theta = 1.2$ [degrees]. All 16 endmember spectra were then used in the final unmixing of the image cube.

The spectra for 12 of these have been plotted (in red) in Figure 1; for those cases where a positive identification has been made, a plot of the corresponding PIMA spectrum is superimposed (in black). In these plots, the wavelength (x-axis) ranges from 2050 nm to 2320 nm while the amplitudes of the reflectances have been individually scaled between the minima and the maxima for comparison. To the immediate right of each spectral plot is the scatter plot in PC space, projected onto the PC plane in which the pixels, which have contributed to that particular endmember spectrum, lie on the periphery of the cloud; these pixels are printed in red, and in addition, a red box has been superimposed to highlight the location of the endmember pixels. To the right of this is the corresponding abundance (fraction) image resulting from the unmixing. In these abundance images, yellow indicates the lowest fractions starting at 0, red, magenta, and blue indicate increasingly higher values, with cyan representing the highest fractions ending at 1. The colour scale contained in Figure 1 applies to all the EM abundance images.

All but one of the 16 endmembers are considered to be significant, i.e. not simply a 'noise' spectrum. This determination is based upon the spatial distribution of the endmember abundance images. If the abundance map for a given endmember exhibits a coherent spatial pattern, particularly if correlated with observed features in the image, then it is concluded that this endmember is related to a particular target species. If, on the other hand, the abundance image is more or less uniformly distributed throughout the image with no correlation to the physical features in the image, then it is concluded that the endmember is primarily noise. Under this criterion, only one of the 16 endmember spectra could have been discarded as noise; this endmember is not displayed here. Also not shown in Figure 1 is the spectrum usually called the 'shadow' endmember, as well as a spectrum that closely resembles endmember 1, and a spectrum that is similar to the alunite spectrum.

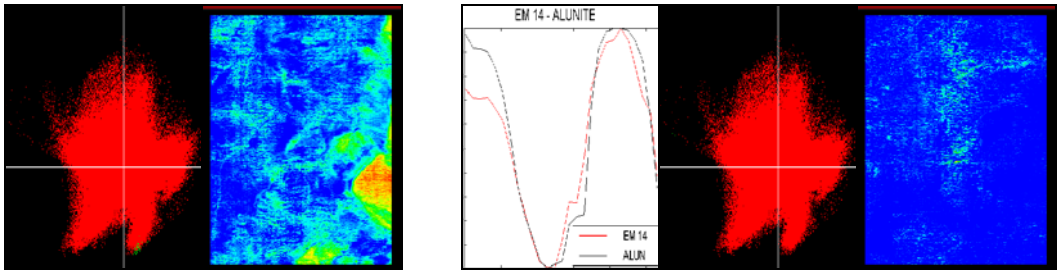
While it is not a part of the IEA procedure, a PC transformation has been performed so that the IEA method can be compared to the common technique in which one first performs a PC transformation, and then extracts the endmembers manually by selecting those vectors which lie at the extremities of the scatter plot in the multi-dimensional PC space. For all but two cases, one of which is endmember 13, the pixels corresponding to the automatically extracted endmembers have been found at or near the extremities of one of the many possible projections onto PC planes. In many of these cases it appears that the selected vectors are contained in an extremity that would be better viewed in a plane other than one orthogonal to one of the PC axes. It is believed that in the remaining two cases, the corresponding extremities are simply not readily viewed in a projection onto one of the orthogonal PC planes; hence, the scatter plot for endmember 13 is not shown in Figure 1. It is expected that, were one to project the data cloud onto other planes in PC space, then one would find all of the automatically extracted endmember pixels located in the extremities of the scatter plot. It is noted, however, that it would have been difficult to find some of these endmembers using the manual technique.


EM 1
EM 1
EM 3 - ALUNITE
EM 3
ALUN
EM 4 - KAOLINITE
EM 4
KAOL
EM 5 - BUDDINGTONITE
EM 5
BUDD
EM 6
EM 6
EM 8
EM 8
EM 9
EM 9
EM 10 - KAOLINITE
EM 10
KAOL
EM 11
EM 11
EM 12 - OPAL CUP
EM 12
OPAL
EM 13 - KAOLINITE
EM 13
KAOL
EM 14 - ALUNITE
EM 14
ALUN

Not

0 Available

Figure 1. Endmember (EM) spectra compared to PIMA spectra of ground samples, EM pixel locations in principal component space, and the corresponding endmember abundance images.



2.3 ANALYSIS

For the 7 cases where the endmember spectra have been identified and the PIMA spectra overlaid, the matches are very good. Endmembers 3 and 14 have been identified as alunite; endmembers 4, 10, and 13 are kaolinite; endmember 5 is buddingtonite; and endmember 12 is opal. Endmembers 1, 7 (not shown in Figure 1), and 11, as yet unidentified, are located in areas that are covered with material excavated from an adit or mine shaft. Endmember 6 may be primarily silica. Endmember 8 is located in a relatively undisturbed area that is known to contain a high percentage of 'desert varnish' coated rocks. Although endmember 9 might be mistaken for dickite, ground samples indicate that it is a mixture of alunite and kaolinite, with one mineral embedded in a matrix of the other. Endmember 16, (not shown in Figure 1) resembles alunite, and may be a mixture of alunite and something else.

For those endmembers which represent mixtures of minerals, one confirmed by ground sampling, the other not confirmed, it is noted that these cannot be linearly mixed, otherwise they would not have appeared as endmembers. In the case that is confirmed by ground sampling, one mineral is mixed in the other with inclusions as small as 1 mm. It is interesting to speculate that, in such cases, light has been transmitted through one mineral into the other and then scattered toward the sensor. This would constitute true non-linear mixing, which could be approximated by a model in which the sample in question is considered to consist of a layer of one mineral superimposed on a layer of the other. The fraction of light transmitted by each mineral 'layer' is described by Beer's law, which states that the transmitted fraction is equal to the exponential of the product of the attenuation coefficient and the thickness of the layer. Thus, the resulting reflectance spectrum is proportional to the exponential of the sum of the products of the attenuation coefficients and the corresponding effective thicknesses of the individual mineral layers. This is a topic for further investigation.

3.0 SUMMARY

The Iterative Error Analysis technique for extracting endmember spectra from imaging spectrometer data cubes has been described. It has been demonstrated by applying it to a 30 band cube of imagery acquired with SFSI near Cuprite, Nevada . The 30 spectral bands cover the range from 2050 nm to 2350 nm which contains spectral absorption features diagnostic for clay minerals. Sixteen endmembers were requested of the IEA processing module; all but one of the resulting 16 have been deduced to be physically significant, while the 16^{th} is considered to be noise. Seven have been positively identified on the basis of a comparison with ground sample PIMA spectra: 2 are alunite, 3 are kaolinite, one is buddingtonite, and one is opal. It has also been demonstrated that the endmembers found by the IEA method include all those one might expect an astute operator to find, and in addition a few that would have been difficult to find manually.

4.0 ACKNOWLEDGEMENTS

We gratefully acknowledge the cooperation of Borstad Associates Ltd. in making available these data acquired by them on the Nevada flights. The Nevada mission was funded by the following companies: Borstad Associates Ltd., Spectral International Inc., Barrick Gold Corporation, BHP Minerals Canada Ltd., Cameco Corporation, Cominco Ltd., CRA Exploration Pty. Ltd., Homestake Mining Company, Newmont Gold Company, North Mining Inc., Placer Dome Exploration, and Western Mining Corporation. Data used in this paper have been provided courtesy of these companies.

5.0 REFERENCES

Adams, J.B., Smith, M.O., and Johnson, P.E. , "Spectral Mixture Modelling: A New Analysis of Rock and Soil Types at the Viking Lander Site," *J. Geophysical Research*, Vol. 91, pp. 8098-8112, 1986.

Anger, C.D., Achal, S., Ivanco, T. Mah, S., Price, R., and Busler, J. , "Extended Operational Capabilities of *casi*." In *Proceedings of the Second International Airborne Remote Sensing Conference*, San Francisco, California, pp. 124-133, 1996.

Boardman, J.W. , "Analysis, Understanding and Visualization of Hyperspectral Data as Convex Sets in *n*-Space." In *Proceedings of the International SPIE Symposium on Imaging Spectrometry*, SPIE Vol. 2480, Orlando, Florida, pp.23-36, 1995.

Neville, R.A., Rowlands, N., Marois, R., and Powell, I. , "SFSI: Canada's First Airborne SWIR Imaging Spectrometer," *Canadian Journal of Remote Sensing*, Vol. 21, No. 3, pp. 328-336, 1995.

Neville, R.A., Nadeau, C., Levesque, J., Szeredi, T., Staenz, K., Hauff, P., and Borstad, G.A. , "Hyperspectral Imagery for Mineral Exploration: Comparison of Data from Two Airborne Sensors." In *Proceedings of the International SPIE Symposium on Imaging Spectrometry*, SPIE Vol. 3438, San Diego, California, pp. 74-82, 1998.

Staenz, K., and D.J. Williams, 1997, "Retrieval of Surface Reflectance from Hyperspectral Data Using a Look-Up Table Approach," *Canadian Journal of Remote Sensing*, Vol.23, No.4, pp. 354-368, 1997.

Szeredi, T., Staenz, K., and Neville, R.A. , "Automatic Endmember Selection: Part I Theory," Submitted for publication in *Remote Sensing of Environment*, 1999.

Vane, G., Green, R.O., Chrien, T.G., Enmark, H.T., Hansen, E.G., and Porter, W.M. , "The Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)," *Remote Sensing of Environment*, Vol. 44, pp. 127-143, 1993.