N95- 23848

AUTOMATING SPECTRAL UNMIXING OF AVIRIS DATA USING CONVEX GEOMETRY CONCEPTS

Joseph W. Boardman

Center for the Study of Earth from Space
Cooperative Institute for Research in the Environmental Sciences
Campus Box 449 University of Colorado
Boulder, CO 80309-0449

1. INTRODUCTION

1.1. Spectral Mixture Analysis and Its Current Limitation

Spectral mixture analysis, or unmixing, has proven to be a useful tool in the semi-quantitative interpretation of AVIRIS data (Boardman and Goetz, 1991 and references therein). Using a linear mixing model and a set of hypothesized endmember spectra, unmixing seeks to estimate the fractional abundance patterns of the various materials occurring within the imaged area. However, the validity and accuracy of the unmixing rest heavily on the "user-supplied" set of endmember spectra. Current methods for endmember determination are the weak link in the unmixing chain.

1.2. Goals of Automated Unmixing and Its Promise

Automated unmixing seeks to estimate the number of endmembers, their spectral signatures and abundance patterns, using only the mixed data and a physical model. It should be an objective and repeatable process that uses no ground-based information. Such a method promises to take full advantage of the wealth of information currently "locked-up" in AVIRIS data sets. The method outlined here seeks to keep this promise by using spectral mixing inherent in AVIRIS data to its **advantage**. These ideas are an extension and follow-on to those first proposed by Craig (1990).

2. A GEOMETRIC VIEW OF SPECTRAL MIXING AND UNMIXING

2.1. Spectra are Points in an *n*-Dimensional Scatterplot

Spectra can be thought of as points in an n-dimensional scatterplot, where n is the number of bands. The coordinates of the points in n-space consist of n-tuples of values that are simply the spectral radiance or reflectance values in each band for a given pixel. An understanding of this concept of "spectral space" is crucial for the following discussion. The distribution of these points in n-space can be used to estimate the number of spectral endmembers and their pure spectral signatures. Although collected in 220 bands, the inherent dimensionality of AVIRIS data is typically much lower, in the range of 3 to 10. This degeneracy is illustrated by the high correlation among bands.

2.2 Convex Geometry : Facets; Faces; Vertices; Hulls; and Simplices in *n*-D

Convex geometry deals with the geometry of convex sets (Valentine, 1964) and, despite it richness and depth, is within the grasp of anyone with a vivid imagination and geometric intuition. Many useful applications have been developed (Lay, 1982). A convex set in n-dimensions is defined as a set of points that are linear combinations of some set of points, where the weights are all positive and sum to unity. This is also the exact definition of the linear spectral mixture model.

Some of the terminology used in convex geometry is illustrated in a 2-d case in Figure 1. All the points are interior to the triangle, since they are positive, unit-sum, linear combinations of the three corners. A body made up of $n+1$ points is the simplest possible, that has some interior, and is called a **simplex**. In this 2-d case it is a triangle. Sub-elements of a convex body are its **faces**. The *(n-1)*-d faces of a simplex are called **facets**. The *0*-d faces are **vertices**. The exterior surface of a set of points, made up of adjacent facets, is its **convex hull**. It consists of the facets that would get "painted" if you "rolled" the point set on a *(n-1)*-d "ink pad". In *3*-d the simplex is a tetrahedron and it has four vertices and four triangle facets. These concepts and definitions generalize to *n* dimensions, despite being harder to visualize. In *n*-d the simplex has *n+1* vertices and *n+1* facets of dimension *n-1*.

2.3. Unmixing as a Convex Geometry Problem

Automated spectral unmixing is a convex geometry problem. We are given the scatterplot of points. Determining its inherent dimensionality tells us the number of mixing endmembers. There are *n+1* endmembers if the data is inherently *n*-dimensional, assuming the endmembers are spectrally distinct in terms of the observing instrument. Estimating the spectral signatures of the mixing endmembers is done by finding the "best-fitting" simplex that contains the scatterplot. The vertices of this simplex represent the *n+1* mixing endmembers. Estimation of the abundances of each endmember, at each pixel, corresponds to a simple transformation of the data to barycentric coordinates. Since the simplex contains all the data points, the derived fractions will be positive and sum to unity, as desired.

3. PRACTICAL METHODS FOR APPLICATION

Applying this endmember derivation approach to AVIRIS data is done in several steps. First the observed radiance data is reduced to apparent reflectance using an atmospheric and solar model, ATREM (Gao and Goetz, 1993). The next step is to determine the data dimensionality and to separate signal from noise. A modified version of the MNF-transform (Green et al., 1988) is used, in place of standard PC analysis, to address the noise properties of AVIRIS data. This involves: estimation of the noise covariance matrix; a rotation and scaling of the data to make the noise isotropic with unit variance in all bands; and a subsequent eigen-analysis of this transformed data. In this MNF-space the number of valid dimensions can be determined by joint inspection of the images and the final eigenvalues. Usually a small number of dimensions explain almost all the signal, with the complementary bands associated with salt-and-pepper noise images and unit eigenvalues. A basis for the signal subspace is calculated and the data are projected onto it, fixing its dimension. The output of this process is apparent reflectance data, projected onto a spanning subspace of minimum dimension, with most of the noise removed.

The estimates of the spectral endmembers are determined by finding the "best-fitting" simplex that contains the projected data. Craig (1990) proposes the "smallest" simplex containing the data as a proxy for the "best" simplex. Numerically, he seeks to maximize a determinant subject to inequality constraints. Geometrically, this shrinks the simplex but keeps all the data inside. Other methods for "best" simplex determination are the subject of current research. Orientations and positions of the *n+1* facets of the "best" simplex may also be determined through analysis of the facets of the convex hull of the projected data. Sets of pixels that are void in one or more endmembers actually map the faces of the desired simplex. Recognizing these "flats" is one way to determine the "best" simplex". Once the best simplex is determined, the endmember estimates are given by its vertices. These endmembers must give positive fractions for abundances that sum to unity for every pixel, since the simplex contains the data. Finally, the data are unmixed by converting to barycentric coordinates and the endmember spectra are projected back to the original band-space to derive their full-resolution spectra.

The validity of the process can be assessed by its outputs. Automated unmixing delivers estimates of the endmembers and estimates of the spatial patterns of fractional abundance of these endmember materials. The endmember spectra should be reasonable and be identified with real Earth materials and "shade". The corresponding spatial abundance images, interpreted together with the spectra, provide another performance check. The spatial information in the abundance images is useful in interpreting and naming the derived endmember spectra, capitalizing on the unique dual spectral/spatial nature of AVIRIS data. Ideally, each endmember spectrum and corresponding abundance image will be reconciled with a named material, perhaps through the use of a spectral library and expert system. The derived spectra have used all bands of all pixels and thus have a much higher signal-to-noise ratio than individual spectra. Furthermore, there are quite few of them by comparison to the raw spectra. Time-intensive procedures can be used to analyze this small set of low-noise, derived spectra, with the unmixing results allowing interpretation of the full set of observed spectra.

4. EXAMPLE APPLICATION TO AVIRIS DATA

Convex geometry unmixing has been applied to many AVIRIS scenes, yielding promising results in both vegetation and geological study areas. Two examples will be shown at the workshop, Jasper Ridge and North Grapevine Mountains.

4.1. Jasper Ridge Example

In this example, 172 bands of 1992 AVIRIS data yielded a *4*-d convex data set. The five endmember spectra and their corresponding abundance images are shown in Slide 2. In this case the interpretation of the endmembers is clear. The five derived endmembers are : shade; water; soil; green vegetation; and dry vegetation. The shapes of the spectra are reasonable and seem to represent "pure" materials, even in cases like shade where no image pixel was nearly pure. This illustrates the ability of the method to project beyond the observed mixed data to estimate endmembers. Pure pixels are not required.

4.2. North Grapevine Mountains Example

In this example, 45 bands of the SWIR portion of 1989 AVIRIS data were used. These data were reduced to approximate apparent reflectance by the empirical line method (Kruse et al., 1993). Using this restricted set of bands, four endmembers were identified. The best-fit simplex and the data are shown in Figure 2, along with plots of two of the endmembers. The four endmember identified are : shade; featureless materials; carbonate; and sericite. The maps of sericite and carbonate abundance agree well with previous studies (Kruse et al., 1993) and the derived spectra are easily identified.

5. DISCUSSION , CONCLUSIONS AND CAVEATS

Convex geometry can be used successfully to address the most important questions in spectral unmixing, the estimation of the number of the endmembers and their spectra. The method outlined here builds on the work of Craig (1990), treating shade as another unknown, obviating the need for a "dark-point" bias correction. Combining AVIRIS data, ATREM and the convex geometry method, one can derive the number of endmembers, estimates of their pure spectra and maps of their apparent abundance, **using absolutely no ground data.**

The method presented here has limitations and pitfalls. Some are inherent, some can be solved and removed with further research. It assumes a wide range of fractional abundances, so it cannot unmix a single pixel, or a homogeneous scene. The current computer programs are limited to scenes with no more than 8 or so endmembers, because of the "curse of dimensionality" (Craig, 1990). Any material that occurs in a fixed proportion in every pixel is essentially "invisible" to the method. This can cause the other derived endmembers to actually be spectra of mixtures, not pure materials.

6. ACKNOWLEDGMENTS

This work was supported by Alex Goetz at CSES, through NASA grant NAS5-31711. Fred Kruse donated AVIRIS data and helpful advice. The concepts of the method are a direct result of work done at CSIRO, DEG, Sydney, Australia. My understanding and thoughts draw heavily on interaction with Australian colleagues: Jon Huntington; Maurice Craig; and Andy Green. Blame is mine, credit theirs.

7. REFERENCES

Boardman, J. W. and A. F. H. Goetz, 1991, "Sedimentary Facies Analysis Using Imaging Spectrometry : a Geophysical Inverse Problem", in : *Proceedings of the Third AVIRIS Workshop*, JPL Publication 91-28, pp. 4-13.

Craig, M., 1990, "Unsupervised Unmixing of Remotely Sensed Images", in: *Proceedings of the Fifth Australasian Remote Sensing Conference*, Perth, Australia, pp. 324-330.

Gao, B-C, K. B. Heidebrecht and A. F. H. Goetz, 1993, Derivation of Scaled Surface Reflectances from AVIRIS Data, *Remote Sensing of Environment*, vol. 44, no. 2, pp. 165-178.

Green, A. A., M. Berman, P. Switzer and M. D. Craig, 1988, "A Transformation for Ordering Multispectral Data in Terms of Image Quality with Implications for Noise Removal", *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65-74.

Kruse, F. , A. B. Lefkoff and J. B. Dietz, 1993, Expert System-Based Mineral Mapping in Northern Death Valley, California/Nevada, Using the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS), *Remote Sensing of Environment*, vol. 44, no. 2, pp. 309-336.

Lay, S. R., 1982, *Convex Sets and their Applications*, John Wiley and Sons, New York.

Valentine, F. A., 1964, *Convex Sets*, McGraw-Hill Book Company, New York.

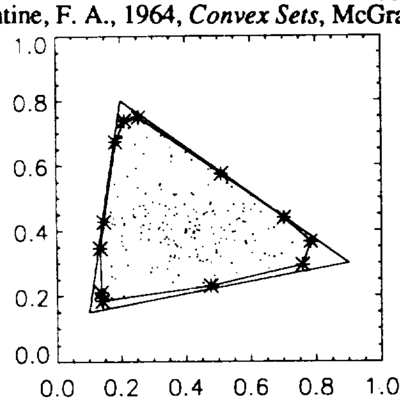

Figure 1. Two hundred points, interior to a triangle illustrate the concepts of convex geometry. The points are all positive, unit-sum combinations of the triangle vertices. Twelve of the points, marked with asterisks, are the vertices of the convex hull. The convex hull has facets that are line segments. All the data are surrounded by the triangle, a 2-d simplex, that has 3 *1*-d facets and 3 *0*-d vertices.

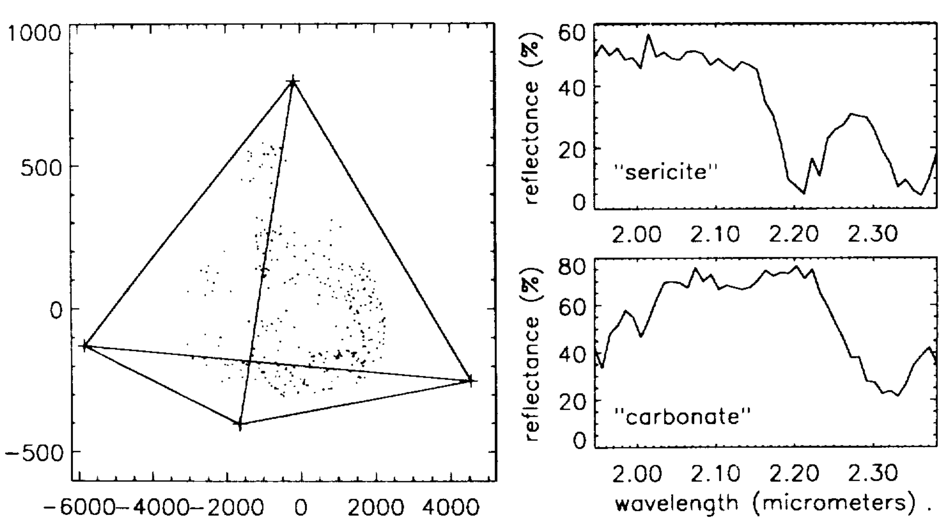

Figure 2. a) Best-fitting simplex surrounds the vertices of the convex hull of the data . Third axis is perpendicular to the paper. b) Two of the derived endmembers, identified as "sericite" and "carbonate". Sericite is the "top" vertex in 2a, carbonate is the "bottom" vertex, shade is to the "right".