Hyperspectral Image Classification Using a Geometrical Model and Stepwise Unmixing

David W. Messinger; Digital Imaging and Remote Sensing Laboratory, Center for Imaging Science, Rochester Institute of Technology; Rochester, NY/USA

Abstract

Hyperspectral imaging produces datasets that have three dimensions: two spatial and one spectral. The spectral dimension typically contains on the order of 200 contiguous bands with spectral resolution on the order of 10 nm, sufficient to perform spectroscopic analysis of the surface. The tradeoff for the high spectral resolution is generally lower spatial resolution resulting in "mixed" pixels. Classification of the imaged surface based on the pixel spectral characteristics into categories such as "urban", "vegetation", "agricultural", and "water" is a common use for such imagery and methods to produce a classification map with a single class assignment per pixel exist. Here, we present a geometrical model of the high-dimensional data in conjunction with a stepwise un-mixing process to produce not a single class map, but a cube of per-pixel class abundances. The method is applied to data taken with the Hyperion sensor on board the EO-1 satellite. End member spectra representing pure materials are first derived from the scene. These are then used in a stepwise regression to find the "best" linear mixture model on a per-pixel basis. Once each pixel is modeled, redundant end member spectra are combined into a subset of classes. The computed mixing fractions are used to "partially classify" each pixel (e.g., 75% class 1 and 25% class 3). Given a classification map where each pixel can be represented by several classes based on the fractions, questions of how to optimally visualize these data arise. Some initial initiatives into this problem are also discussed.

Introduction

Classification of airborne and space-based imagery into well-defined land classes is a common objective of the remote sensing community. This classification is based on the spectral characteristics of the measured pixel radiances. Typical land classes include general categories such as "urban", "vegetation", "agricultural", or "water". Other classification schemes seek to answer questions such as trafficability or arability. Hyperspectral sensors are well-suited for such an application as the spectral content of the retrieved pixel contains unique information about the material through its reflectance properties. Hyperspectral sensors in the reflective portion of the EM spectrum typically have ∼200 spectral channels ranging from the visible to the near infrared (0.35 to 2.4 μm) with a spectral resolution of ∼10 nm. Observations can be complicated, though, by the presence of the intervening atmosphere for airborne and space-based observations, thus complicating the exploitation tasks. As a trade-off for the high spectral resolution of these sensors, the pixel sizes are typically large. Space-based systems have ground sample distances (GSD) on the order of 30 m; airborne systems can have pixel sizes on the order of 2 - 20 m, depending on sensor specifics and collection geometries. Such large pixel sizes result in "mixed" pixels due to the spatial varibility of materials providing few uniform regions at these scales.

Several classification schemes have been developed for both multispectral imagery (containing ∼10 - 20 spectral bands) and hyperspectral imagery[1, 2, 3]. These fall into two categories: supervised and unsupervised. The former method requires the analyst to identify training regions for each class anticipated to be in the image. Each pixel under test is then compared (in a statistical sense) to the pre-defined class training regions to identify the class to which it most likely belongs. Unsupervised methods use little or no intervention on the part of the analyst, instead relying solely on the scene content to classify each pixel. Usually, the number of classes into which the scene is to be classified is pre-determined. A common method is the k-means, or moving means, method[4]. The number of classes is pre-determined, and the means of those classes are initially randomly chosen. All pixels are classified based on a distance measure from the class means, and new means are then computed. The process iterates until there is little or no change in the classifications. All methods have in common the use of the spectral content of the pixel under test.

Here, an unsupervised method is presented that uses scene-derived end member spectra (assumed to represent pure, distinct materials) in a geometrical model of the data space. These end member spectra are then used in an iterative regression method to model each pixel with the most appropriate end members and thus, classify each pixel. The approach uses a linear mixing model[5] combining end member spectra to model each measured pixel spectrum. As a result, the mixing fractions, or abundances, of the particular end member spectra used to model each pixel are reported relating the contents of the pixel to user-identified classes.

This paper is organized in the following manner. First, the algorithm is described, giving details about the identification of the end member spectra from the scene, the regression approach, and the pixel classification method. Next, the test data are presented. Results from the classification scheme are shown and issues regarding the visualization of such results are presented. Finally, a summary of the work concludes the paper.

Algorithm Description
End Member Selection

Hyperspectral imagery can be modeled in a geometric sense by treating every pixel spectrum as a vector in an *n* dimensional space, where *n* is the number of spectral bands in the sensor. Every pixel in the image is assumed to be a linear mixture of pure component spectra associated with the materials in the image. These pure component spectra, called end members, can be derived from the image in several ways. Most simply, a user can visually identify "pure" pixels and extract those spectra. Automated methods exist for end member extraction as well[6]. The method used here is termed MaxD and is

based on the concept that hyperspectral data lie within a convex hull in the *n*-dimensional space[7]. Given this shape of the data cloud, each interior point in the cloud can be modeled by a linear combination of the spectra at the corners of the simplex enclosing the data. These vectors can be shown to span the *n*-dimensional space.

The algorithm uses geometrical projections within the *n*-dimensional space to determine the corners of the simplex. Each pixel spectrum within the image is treated as a vector in the space. The points with the largest and smallest Euclidean magnitudes are identified as the first two end members. All the data are subsequently projected along the vector between these two points. This operation not only places the two end members "on top of" each other, but also retains all corners in the simplex as corners. Now, the data point (in the newly projected space) that is furthest from the point occupied by the first two end members is considered the third end member, and all data are projected along the vector connecting the first two spectra (now a single point) and the new extreme. This process is iterated until as many end members as desired are extracted from the scene. The scheme is automated requiring the user to supply only a stopping point. These end members are then used in the regression scheme, described below, to model each pixel in the scene.

Stepwise Regression

If we assume that each pixel in the scene is a linear combination of the end members representing the pure materials (*i.e.*, classes), then all that remains is to determine "how much" of each end member is in a particular pixel. This is commonly done using a linear mixing model,

$$L(\lambda) \simeq \sum_{i=1}^{N} \alpha_i \hat{L}_i(\lambda). \tag{1}$$

Here, $L(\lambda)$ is the test pixel spectrum to be modeled (here in radiance space, but possibly in reflectance space if the image has been atmospherically-compensated), $\hat{L}_i(\lambda)$ are the end member spectra previously determined, and α_i are the mixing fractions. Given the end members, this model is simple to fit using a least squares regression returning the mixing fractions α_i for all end members in the set.

For high-altitude airborne and space-based hyperspectral imagery with large collection areas, it is not reasonable to assume that *every* end member, or material class, will be present in *every* pixel in the scene. A more physical model of the individual pixel would choose which end members are in that pixel, and exclude all others from the final fitting process. To achieve this, a stepwise regression technique has been implemented using an Analysis of Variance (ANOVA) metric to determine which set of end members best represents each pixel on a per-pixel basis.

In this method, each pixel is modeled first using each end member individually. The single end member that achieves the best fit to the measured pixel spectrum is kept in the model. Then, every two-spectra combination of that first end member and every other end member is used to model the measured pixel spectrum. The best combination of two end members is then kept in the model. This process is continued to include a third end member. However, because the combination of just the second and third end members has not been tested, the first end member is removed from the model and a determination is made whether a better fit is achieved with just end members two and three. At each point in the process, the determi-

nation of whether the model is improved by the addition / subtraction of an end member is made using an ANOVA calculation and an f-test. The process continues testing (almost) all possible combinations of end members. Ultimately, each pixel is modeled with the "best" set of end member spectra from the original set. The Linear Mixture Model (equation 1) is used to fit that pixel and the mixing fractions are determined. For those end members not included in the final model for the test pixel, the mixing fractions are set to zero. The solution for the mixing fractions is partially constrained such that $\alpha_i \geq 0$, but they do not have to sum to one.

Classification Scheme

Once each pixel has been fit with a model of appropriate end member spectra, the computed mixing fractions can be used to classify the pixels. First, however, because the end member extraction scheme is unsupervised, the end members themselves must be collapsed into a subset of classes. As it is difficult to determine *a priori* exactly how many classes are contained in a particular scene, there can be redundancy in the end members selected. For this work, twenty end members were derived from the scene originally, and these were reduced into eight classes. A new class map was derived with the final eight classes and fractions for each class were computed by summing the mixing fractions from the original end member fits appropriately (*e.g.*, if end members 3, 5, & 8 were assigned to class #1, the mixing fractions from those three classes were summed and assigned as the mixing fraction for class #1). The resulting eight classes were assigned names based on visual inspection and scene context of the pixels assigned to that class.

Test Data

Data for this experiment were taken with the Hyperion sensor on board NASA's EO-1 satellite[8]. An image of the Rochester metropolitan area was collected in June of 2004. A subset of the entire collection was extracted for use in this work. The subset is of an agricultural area south of Rochester, and includes some small cities as well as part of a large lake (Conesus Lake, part of the Finger Lakes). The Hyperion sensor has 256 spectral channels ranging from $0.35\,\mu$m to $2.5\,\mu$m. Several channels were removed from the data prior to processing due to sensor noise and atmospheric absorption bands. The final cube contained 145 spectral bands and was 256 samples wide by 500 lines long. With a GSD of 30 m the cube covers an area approximately 15 km long by 7.7 km wide. An RGB image of the scene is shown in Figure 1.

Results

Class names associated to the classification results. Classes 1-8 are associated to Figures 2(a) - (h).

Class #	Class Name
1	agricultural - green
2	agricultural - brown
3	agricultural - other
4	forest
5	urban
6	water
7	soil
8	other

Results from the classification scheme are shown in Figure 2. A gray-scale image is shown for each of the final eight classes, with brightness corresponding to higher mixing fraction for that classes. Class names associated with the eight classes are shown in Table 1. The first three of the eight classes into which the scene has been classified are all "agricultural" classes of different visual color (Figure 2(a) - (c)). The primary distinction is between the "green" agricultural fields and the "brown" fields. The image was taken in June, so the supposition is that the "brown" fields are fallow. The "agricultural - other" represents crops that were of a different visible color. These tended to be brighter and slightly more red in the visible image.

The fourth class (Figure 2(d)) represents the "forest" class. This class is hardly separable from the green agricultural class. This is not unexpected as the green trees are spectrally similar to green crops, particularly at this spatial resolution. The class was named as forest, though, due to the successful classification of the large forest region in the middle left portion of the image. This area is noticeably absent in the mixing fraction maps for all other classes.

Likewise, the "urban" class shown in Figure 2(e) is similar to several areas of the bare soil class (class #7, Figure 2(g)). This class does clearly define the roadways that run through the image, as well as the small urban areas in the upper left corner of the image and at the north end of the lake. Additionally, built-up areas along the shore of the large lake in the lower right corner of the image are highlighted in this class.

Class 6, the "water" class (Figure 2(f)) very dramatically highlights the lake, as well as several smaller bodies of water in the scene. However, because the water represents the "darkest" class in the image, several other dark portions of the image that appear to be agricultural in nature are highlighted with smaller abundance. These could be recently watered fields, which would exhibit the spectral characteristics of water, or simply materials with a lower overall reflectance.

Class 7 is the "soil" class (shown in Figure 2(g)) and, as mentioned above, exhibits much of the same spatial structure (except for the roadways) as the urban class. This class spatially exhibits structures associated to the agricultural fields as well. This is a reflection of the fact that at these spatial resolutions (*i.e.*, pixel sizes of 30 m), it is possible that for some crops we are seeing the soil in between the individual plants, thus creating mixed pixels.

Class 8 is the "other" class (Figure 2(h)). Here it mostly contains spectral anomalies within the scene, as well as sensor artifacts.

Visualization of High-Dimensional Data Products

While it is possible to assign a single class to each pixel in this imagery, and imagery like it, at these spatial resolutions (GSD

(a)   **(b)**   **(c)**

(d)   **(e)**   **(f)**

(g)   **(h)**   **(i)**
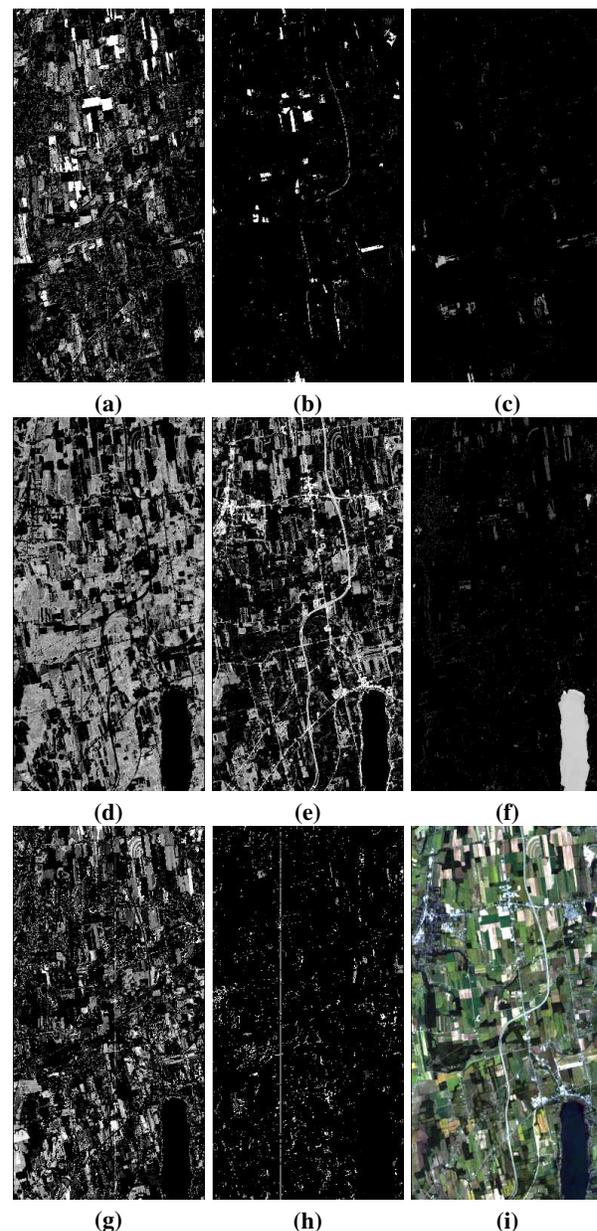
Figure 2. Results from classifying the image into eight classes. Images are for classes 1-8, figures (a) - (h), respectively. Please refer to Table 1 for the class identifications. Figure 1 is shown again for comparison (i).

Figure 1. *RGB image of the Hyperion data used.*

$\simeq 30\,\text{m}$) material continuity is not guaranteed, and a more spatially accurate classification scheme may be necessary. A technique that provides a fractional classification is useful for such circumstances, particularly in transition regions. Unfortunately, this creates a problem for the visualization of the data. Given a single classification per pixel, a simple scheme can be used to demonstrate a large number of classes. When each pixel can have a fractional abundance of each class, it becomes difficult to portray effectively the information on a simple display as the data become more "high-dimensional." Research into this topic is underway[9] attempting to develop methods of visualization based on the information content in the scene. Initial work indicates that methods such as Principle Components Analysis and Independent Components Analysis can be used to reduce the dimensionality of such data products. The reduced-dimension data are then mapped into various color spaces for presentation of the final classification results.

Summary

A method has been presented that uses scene-derived end members, derived from a geometrical model, in a stepwise regression to fit a linear mixing model to each pixel. The mixing fractions resulting from this regression are used as fractional classifications after consolidation of the end member spectra into user-defined classes. This techniques has been applied to a data set collected by the Hyperion sensor on the NASA EO-1 satellite. The area under investigation is an agricultural area and was classified into eight classes. Each pixel has a fractional abundance associated to each class allowing for many class assignments. This presents a significant visualization problem. Several methods are under investigation as to how "best" to represent such high-dimensional data.

References

[1] R.A. Schowengerdt, Remote Sensing Models and Methods for Image Processing, Academic Press, New York, 1997.

[2] J.R. Schott, Remote Sensing: The Image Chain Approach, Oxford University Press, New York, 1999.

[3] J.A. Richards, Remote Sensing Digital Image Analysis: An Introduction, Springer-Verlag, New York, 1999.

[4] R.O. Duda & P.E. Hart, Pattern Classification and Scene Analysis, Wiley, New York, 1973.

[5] N. Keshava & J. Mustard, Spectral Unmixing, IEEE Signal Process. Mag., 19(1), 44 (2002).

[6] M.E. Winter, N-FINDR: An Algorithm for Fast Autonomous Spectral End Member Determination in Hyperspectral Data, Conference on Imaging Spectrometry, Proceedings of SPIE, v. 3753, pg. 266 (1999).

[7] P. Bajorski, E.J. Ientilucci, & J.R. Schott, Comparison of Basis-Vector Selection Methods for Target and Background Subspaces as Applied to Subpixel Target Detection, Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X, Proceedings of SPIE, v. 5425, pg. 97 (2004).

[8] J.S. Pearlman, P.S. Barry, C.C. Segal, J. Shepanski, D. Beiso & S.L. Carman, Hyperion, a Space-Based Imaging Spectrometer, IEEE Trans. on Geoscience and Rem. Sens., 41(6), 1160 (2003).

[9] C. Zhang, D.W. Messinger, E. Montag, Perceptual Display Strategies of Hyperspectral Imagery Based on PCA and ICA, to be published in Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XII, Proceedings of SPIE, v. 6233 (2006)

Author Biography

Dr. Messinger received a Bachelors degree in Physics from Clarkson University and a Ph.D. in Physics from Rensselaer Polytechnic Institute where his Ph.D. thesis project was a study of interstellar infrared polarization due to spectral absorption. He is currently an Associate Scientist in the Digital Imaging and Remote Sensing Laboratory at the Rochester Institute of Technology. There, he leads the Algorithm and Phenomenology Group and performs research in the areas of Hyperspectral Image Exploitation and Simulation.