

On Information Extraction Principles for Hyperspectral Data

by

David Landgrebe
School of Electrical & Computer Engineering
Purdue University
West Lafayette IN 47907-1285
landgreb@ecn.purdue.edu

Abstract

Means for optimally analyzing hyperspectral data has been a topic of study for some years. Our work has specifically focused on this topic since 1986¹. The point of departure for our study has been that of signal theory and the signal processing principles that have grown primarily from the communication sciences area over the last half century. The basic approach has been to seek a more fundamental understanding of high dimensional signal spaces in the context of multispectral remote sensing, and then to use this knowledge to extend the methods of conventional multispectral analysis to the hyperspectral domain in an optimal or near optimal fashion. The purpose of this paper is to outline what has been learned so far in this effort.

The introduction of hyperspectral sensors that produce much more detailed spectral data than those previously, provide much enhanced abilities to extract useful information from the data stream that they produce. In theory, it is possible to discriminate successfully between any specified set of classes of data by increasing the dimensionality of the data far enough. In fact, current hyperspectral data, which may have from a few 10's to several hundred of bands, essentially make this possible. However, it is also the case that this more detailed data requires more sophisticated data analysis procedures if their full potential is to be achieved. Much of what has been learned about the necessary procedures is not particularly intuitive, and indeed, in many cases is counter-intuitive. In this paper, we shall attempt not only to illuminate some of these counter-intuitive aspects, but to point the direction for practical methods to make optimal analysis procedures possible.

Key Words: Hyperspectral Data, Signal Spaces, Hyperspectral Analysis.

1

© 1999 by David Landgrebe. Copying for personal or non-commercial educational use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating other new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from David Landgrebe. Work leading to the material presented here was funded in part by NASA Grants NAGW-925(1986-94), NAG5-3924 (1994-97), and NAG5-3975 (1997-2000). Additional work was done under Army Research Office (MURI) Grant DAAH04-96-1-0444 (1996-1998). This support is gratefully acknowledged.

Introduction and Background

Imagery has formed the basis for acquiring information remotely for many years. Aerial photography has been used extensively for this purpose almost since photography was invented. With the launching of Sputnik, the world's first artificial satellite in 1957, it was natural that the possibility of using spacecraft be considered for this purpose. The first thoughts focused upon monitoring the weather, and the first satellite designed for this purpose was launched on April 1, 1960. It was not long after that attention also turned to land resources in addition to weather.

The first thoughts for land remote sensing from space again focused on imagery. However, this was soon paralleled with multispectral line scan data, and thus the Earth Resources Technology Satellites were configured to have both return beam vidicons for framed images and a multispectral scanner (MSS) for spectral data. Thus, in addition to images as such, the focus was also placed upon the spectral distribution of the energy emanating from each pixel.

The following graph shows a timeline of the development of the multispectral concept, in terms of key sensor system milestones.

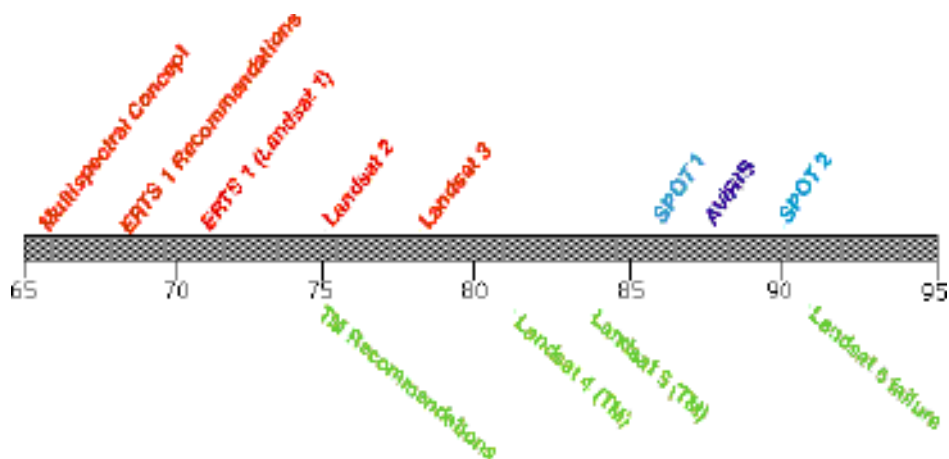


Figure 1. A timeline on the development of land remote sensing satellite systems.

The concept of using primarily the spectral distribution of energy from each pixel as the basis for identification of the pixel contents, as opposed to inherently image oriented properties became crystallized in the mid 1960's and became known as the multispectral concept. Research on it quickly led to specifications for the MSS to be on board the first generation of a land oriented series of satellites by 1968. The series, initially known as the Earth Resources Technology Satellites, ERTS, was soon renamed Landsat, with three sequentially launched spacecraft in the 1970's. In 1975, specification for a second generation system, to be known as Thematic Mapper, were arrived at and form the basis for the Landsat system still flying today.

Progress toward practical means for analyzing multispectral data of land surface areas was very rapid during the 1960's and 1970's. The primary limiting factor during that time was the rather crude spectral characterization of the reflected and emitted energy from the land surface subject

material (3 to 7 spectral bands and 6 to 8 bit precision). The coming of hyperspectral sensor technology over the last few years has largely removed that limitation. The AVIRIS system, with its 210 bands in the 0.4-2.4 μm region was a groundbreaking early example. This resulted in the primary limiting factor to deriving useful scientific and practical information from such data becoming the deriving of suitable means for analyzing the much more complex hyperspectral data. The potential for information extraction is so much greater for hyperspectral data, but the price for achieving this potential is in the need for greater sophistication in the analysis procedures required.

It appears that the progress achieved over the last decade or so on multispectral and hyperspectral analysis technology in the field at large has not been as great as might be expected, given the substantial effort by a large variety of workers. This is what stimulated the beginning of an approach to the problem from an alternate, more fundamental point of view. The intent is to come to understand the first principles controlling the extraction of information from such high dimensional data and to see if a more effective approach could be found to the problem.

Currently Prominent Analysis Methods

Briefly stated, the large body of what has gone on in the field at large has been to try to correct the data of a new data set for the confounding observational and measurement factors that arise in the data collection process. This has turned out to be a quite daunting and perhaps never-ending challenge. Further, the central idea has apparently been to do the correction so that the new data set can be related to existing spectral reflectance curves for each material of interest. Some of the limitations of this approach are that,

1. Rarely can the observational parameters, e.g. solar illuminations factors, atmospheric effects, non-lambertian surface characteristics, etc. be measured to adequate precision on a pixel by pixel basis as would be required to make an adequately precise data adjustment, and
2. The highly dynamic nature of the reflectivity of Earth surface classes is easy to underestimate, such that the existence of such standard spectral responses (spectral signatures) that are adequately stable from time to time and place to place is called into question.

Further, such approaches tend to be based on single spectral curves to characterize a given target material. This ignores significant aspects of spectral responses that have been shown to be quite diagnostic in nature.

The Basis for a Hyperspectral Analysis

Thus, rather than approaching the problem by trying to adjust a new data set to previous standards, one might focus on learning how to model the classes of interest within the original data set itself to adequate precision using information that is likely to be available to an analyst at the time the analyst begins the analysis process.

This leads one to ask just what about multispectral image data is information-bearing? The matter of how spectral variations are represented mathematically and conceptually is an important first step in answering this question and in defining how the extraction of desired information should proceed. There have been three principal ways in which multispectral data is represented quantitatively and visualized. See the Figure 2.

- In image form, i. e., pixels displayed in geometric relationship to one another,
- As spectra, i. e., variations within pixels as a function of wavelength,
- In feature space, i. e., pixels displayed as points in an N-dimensional space.

We will refer to these three as image space, spectral space and feature space.

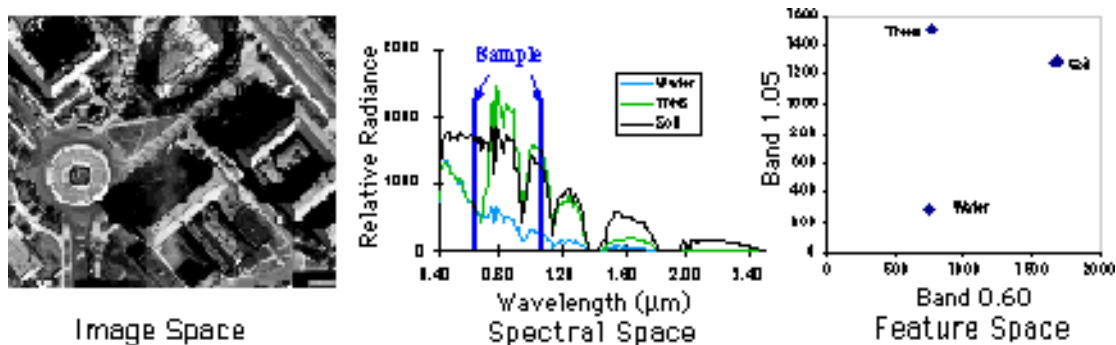


Figure 2. Conceptual Examples of the three ways in which multispectral data may be represented. With the multispectral concept, each pixel in Image Space would be represented as a single curve in Spectral Space. Then, if a spectral curve is sampled at, for example, the two wavelengths indicated, the result can be plotted as a point in Feature Space as shown. Sampling a spectrum at an adequate number of wavelengths can result in a spectral response being fully represented, but as a point in a high dimensional space.

Image Space. Though the image form is perhaps the first form one thinks of when first considering remote sensing as a source of information, its principal value has been somewhat ancillary to the central question of deriving thematic information from the data. Data in image form serve as the human/data interface in that image space helps the user to make the connection between individual pixels in the data and areas on the ground, and the surface cover classes they represent. It also supports area mensuration activities usually associated with remote sensing techniques. For this reason, it becomes very important as to how accurately the true geometry of the scene is portrayed in the data. However, it is the latter two of the three means for representing data that have been the point of departure for most multispectral data analysis techniques.

Spectral Space. Many analysis algorithms that appear in the literature begin with a representation of a response function as a function of wavelength. Early in the work, the term "spectral matching" was often used, implying that the approach was to compare an unknown spectrum with a series of pre-labeled spectra to determine a match, and thereby to identify the unknown. This line of thinking has led, at various times, to attempts to construct a "signature bank," a dictionary of candidate spectra whose identity had been pre-established. Though an attractive and straightforward idea, spectral matching and signature banks have not proven to be very powerful in terms of their ability to extract information in a robust and practical sense.

A second example of the direct use of spectral space is the "imaging spectrometer" concept, whereby identifiable features within a spectral response function, such as absorption bands due to resonances at the molecular level, can be used to identify a material associated with a given spectrum. This approach, arising from the concepts of chemical spectroscopy, which has long been used in the laboratory for molecular identification, is perhaps one of the most fundamentally cause/effect based approaches to multispectral analysis, however, it, too, has its limitations in practical circumstances.

Feature Space. The third basis for data representation also begins with a spectral focus, i.e., that energy radiance or reflectance vs. wavelength contains the desired information, but it is less related to pictures or graphs. It began by noting that the function of the sensor system inherently is to sample the continuous function of emitted and reflected energy vs. wavelength and convert it to a set of measurements associated with a pixel which constitute a vector, i.e., a point in an N-dimensional vector space. This conversion of the information from a *continuous* function of wavelength to a *discrete point* in a vector space is not only inherent in the operation of a multispectral sensor, it is very convenient if the data are to be analyzed by a machine-implemented algorithm. It, too, is quite fundamentally based, being one of the most basic concepts of signal theory. Further, it is a convenient form if a more general form of feature extraction is to precede the analysis step, itself. As will be seen below, of the three data representations, the feature space provides the most powerful one from the standpoint of information extraction.

Next, consider how multispectral data typically appears in feature space. We will use a particularly simple situation to illustrate this. Figure 3 shows a scatter plot of two bands of Landsat Thematic Mapper data for an agricultural area. The area involved covers only two agricultural fields, one containing corn and the other soybeans. One sees from this graph that the separability of these two classes is not apparent from the scatter plot. The different crop responses do not manifest themselves as relatively distinct clusters. Rather, the data distributes itself more or less in a continuum over this space. This is typical of multispectral data, and indicates that the characteristics that allow discrimination between classes are more subtle than such straightforward examination would permit. In fact, a maximum likelihood classification of these data yields a 79.8% accuracy for the two classes corn and soybeans. If all 7 bands of the data are used, this accuracy rises to 100%. Viewed in image space these two areas appear essentially identical; in spectral space, where only first order spectral variations are apparent, they appear heavily overlapped, and yet, tested quantitatively in feature space, they are shown to be quite separable.

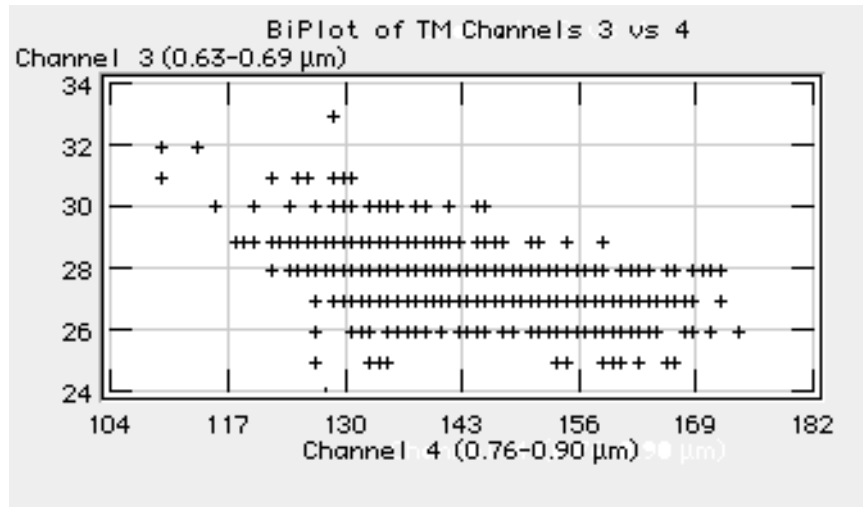


Figure 3. Scatter plot for two channels of 750 Landsat TM pixels from the classes soybeans and corn.

Types of Problems

The types of uses to be made of information derived from multispectral aircraft and spacecraft data vary widely. From the standpoint of analysis technology, an incomplete sampling of the possible categories might be,

- Thematic Maps – Exhaustive Classes. By this is meant a mapping of the original data vectors into an exhaustive list of user-defined categories.
- Thematic Maps – Non-exhaustive Classes. As the complexity of multispectral data has grown, and thus its potential has increased, the number of classes that can be discriminated between has also increased. Many times, the user is only interested in a small number of classes in a scene. The rest may be regarded as background.
- Target Location. One may be interested only in determining if a given material is present in a scene and if so, where.

Each of these requires different, though related methods of data analysis. For reasons of space, we shall focus on the first of these three.

Characteristics of High Dimensional Spaces.

As stated above, while much information can be derived by human observation in image space or spectral space, feature space is most useful for quantitative (machine) analysis. Thus, our focus will be on the feature space representation. Much fundamental knowledge has been acquired about hyperspectral data from this perspective in recent years². The work has been approached

² See, for example, Chulhee Lee and David A. Landgrebe, "Analyzing High Dimensional Multispectral Data," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 31, No. 4, pp 792-800, July, 1993, and Jimenez, Luis, and David Landgrebe, "Supervised Classification in High Dimensional Space: Geometrical, Statistical, and Asymptotical Properties of Multivariate Data," *IEEE Transactions on System, Man, and*

from the standpoint of signal theory as studied in signal processing engineering, and revolves around viewing the data of each pixel as a point in an N-dimensional signal space, where N initially corresponds to the number of bands in the data. This signal space is referred to as feature space because as processing proceeds, linear transformations may be carried out on the data, turning the dimensions of the space into more focused spectral features which can be used in discriminating between classes of interest.

Example specific characteristics of high dimensional feature spaces that are especially relevant to the task at hand are³,

- The geometry of high dimensional spaces is very different from conventional three-dimensional geometry. The volume available in such higher dimensional spaces is very great such that it is likely that any data set to be analyzed will be sparse, not functioning as a dense cloud, but rather as a scattering of points distributed sparsely over a region of the space. As a practical matter, the positive impact of this characteristic is that, since pixels are rarely precisely coincident in such a high volume space, any subclass of pixels can be discriminated from any other subclass if the subclasses can be quantitatively described to adequate precision. For example, if one has 10 bit data in 100 bands, there would be a possibility of data in $1024^{100} = 10^{300}$ discrete cell locations in the space. This number is so large that, even for a data set of 10^6 pixels, the probability of even two pixels occupying the same cell is almost vanishingly small. Thus, since there is no overlap of pixels, the possibility exists, at least theoretically, for any two classes to be separable.
- As a direct impact of this possibility, attention must be focused upon the precision and completeness of the quantitative specification of the classes. Use of a single spectral curve or the average of a number of such curves, while it may be useful in some cases, is clearly limiting in terms of ultimate performance. Higher order statistics of a class, providing information about the shape of a class distribution, takes on added significance as the feature space dimensionality increases. For example, one of the common ways to estimate classification accuracy *a priori* is by use of a statistical distance measure such as the Bhattacharyya Distance. In the Gaussian case, this is given by

$$B = \frac{1}{8} [\mu_f - \mu_g]^T \frac{\Sigma_f + \Sigma_g}{2}^{-1} [\mu_f - \mu_g] + \frac{1}{2} \ln \frac{\left| \frac{1}{2} [\Sigma_f + \Sigma_g] \right|}{\sqrt{|\Sigma_f| |\Sigma_g|}}$$

where μ_i is the mean of class i and Σ_i is its covariance matrix. Note that the first term on the right quantifies the class separation due to the difference in mean values, while

Cybernetics, Volume 28 Part C Number 1, Feb. 1998. Copies of these and other works are available for downloading from

<http://dynamo.ecn.purdue.edu/~landgreb/publications.html>

³ Additional details of these factors may be found in David Landgrebe, "Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data," Chapter 1 of *Information Processing for Remote Sensing*, edited by C. H. Chen, published by the World Scientific Publishing Co., Inc., 1060 Main Street, River Edge, NJ 07661, USA (Spring, 1999).

the second term quantifies the separation due to covariance difference. From this it is apparent that two classes could have exactly the same mean values, but still be quite separable.

- It is characteristic of the remote sensing problem that one is extrapolating from a very limited knowledge of a scene to a detailed, often pixel by pixel mapping of the themes of interest in the scene. This, then, forms the basis of one of the key problems that must be solved in devising a practical and effective analysis process: what form of preliminary information will the analyst have about the scene at the outset, and how can this preliminary information be turned into an adequately precise quantitative description of the classes one desires to identify.
- Other characteristics of high dimensional feature spaces which are especially relevant have been identified. It is known, for example, that because of the high volume of such spaces and the fact that they are thus mostly empty, the significant structure of any specific data set of interest lies in a lower dimensional subspace, but one that varies from data set to data set. This led to a search for good feature extraction schemes, linear transformations that find the subspace in which the classes of interest can best be separated. Several of them have been identified, with families of characteristics that reasonably well span the space of circumstances likely to be encountered. In addition to the well-known discriminant analysis algorithm, these include decision boundary feature extraction^{4,5,6} and projection pursuit.⁷
- The most logical way of quantitatively defining the classes of interest in a given analysis situation is via design or training samples drawn directly from the data set to be analyzed. Doing so, in effect, normalizes for sensor and observation variables that exist in the scene at the time of observation. This fact substantially reduces the need to make data adjustments based upon individual sensor and scene conditions, thus significantly reducing the overall complexity of the processing, making the analysis more robust and easier for the user to carry out.
- As the dimensionality of the feature space increases, the number of such samples needed to adequately characterize the classes rises very rapidly. Again, because the number of such pre-labeled samples is likely to be small compared to the need, the

⁴ Chulhee Lee and David A. Landgrebe, "Feature Extraction Based On Decision Boundaries," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 4, April 1993, pp 388-400.

⁵ Chulhee Lee and David A. Landgrebe, "Decision Boundary Feature Selection for Non-Parametric Classification," *IEEE Transactions on System, Man, and Cybernetics*, Vol. 23, No. 2, March/April, 1993, pp 433-444.

⁶ Chulhee Lee and David A. Landgrebe, "Decision Boundary Feature Extraction for Neural Networks," *IEEE Transactions on Neural Networks*, Vol. 8, No. 1, pp 75-83, January 1997.

⁷ Luis O. Jimenez, "High Dimensional Feature Reduction Via Projection Pursuit," PhD Thesis, Purdue University, May 1996, also available as Luis O. Jimenez and David Landgrebe, "High Dimensional Feature Reduction Via Projection Pursuit," School of Electrical & Computer Engineering Technical Report TR-ECE 96-5, April 1996.

additional factor of the classification algorithm complexity comes into play. It has been found that simpler algorithms outperform more complex ones when the number of samples available is too limited. An algorithm to define the correct degree of algorithm complexity, called LOOC⁸, has been defined to assist with this process, and an algorithm to mitigate the small sample size effects has been created⁹.

A Concept for Data Analysis

Taking into account these factors, what would be a maximally effective hyperspectral data analysis procedure? The procedure needs to take into account the following.

- The process needs to account for more than just first order spectral variations. Higher order variations as conveyed, for example, by the covariance matrix often carry a substantial or even sometimes a majority of the ability to discriminate between classes.
- The process should be conducted in a feature subspace that is determined based upon the specific quantitatively defined classes desired and the data set to be analyzed. In other words the feature space to be used should be as situation-specific as possible and at the optimum dimensionality for the problem at hand. Too many features can be just as bad as too few.
- The final classification process itself needs to be based on the original data itself rather than a processed, corrected version of it. Though they often can greatly aid in human perceptions as to circumstances, data corrections algorithms are usually designed to adjust for first order effects only, and have unknown and even deleterious impact or higher order information-bearing variations in the data.

These factors suggest a procedure as diagramed in Figure 4. Key steps are the feature extraction step, which has become really of fundamental importance with the increase in dimensionality of hyperspectral data, and the circumstances of deriving adequately precise class descriptions. We shall elaborate briefly via examples.

⁸ Hoffbeck, Joseph P. and David A. Landgrebe, "Covariance Matrix Estimation and Classification with Limited Training Data," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, no. 7, pp. 763-767, July 1996.

⁹ Behzad M. Shahshahani and David A. Landgrebe, "The Effect of Unlabeled Samples in Reducing the Small Sample Size Problem and Mitigating the Hughes Phenomenon," *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 32, No. 5, pp 1087-1095, September 1994.

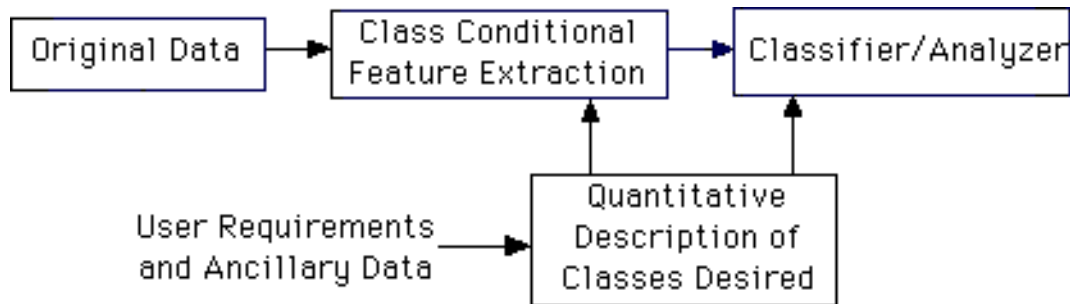


Figure 4. The sequence of analysis steps.

A Feature Extraction Example

Earlier several feature extraction algorithms were listed. Following is a brief illustration of how the feature extraction concept worked in a problem to discriminate between minerals of geologic interest^{10,11}. In this case, it was known that the minerals involved have specific, narrow band absorption features in the 2 μ m region. The four graphs in Figure 5 show these features¹². Hyperspectral data in 210 bands from 0.4 to 2.4 μ m was gathered using the AVIRIS system sensor over a potential mining site of interest in Nevada, with the desire to map these four minerals as expressed on the surface.

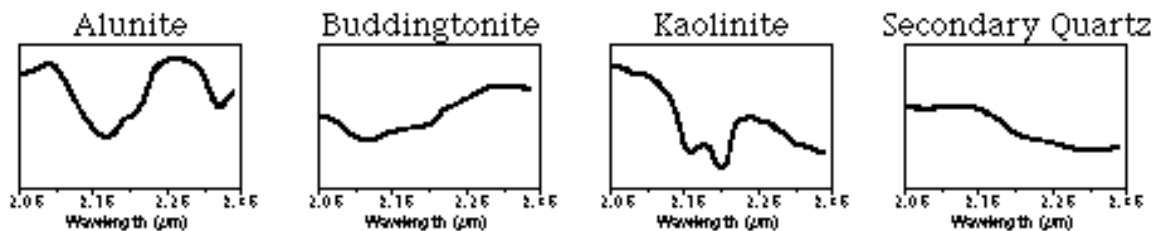


Figure 5. Absorption features for four minerals.

¹⁰ Hoffbeck, Joseph P. and David A. Landgrebe, "Classification of Remote Sensing Images having High Spectral Resolution," *Remote Sensing of Environment*, Vol. 57, No. 3, pp 119-126, September 1996.

¹¹ Joseph P. Hoffbeck, "Classification of High Dimensional Multispectral Data," PhD Thesis, Purdue University, May 1995, also available as Joseph Hoffbeck and David A. Landgrebe, "Classification of High Dimensional Multispectral data," School of Electrical Engineering Technical Report TR-EE-95-4, May, 1995.

¹² 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.

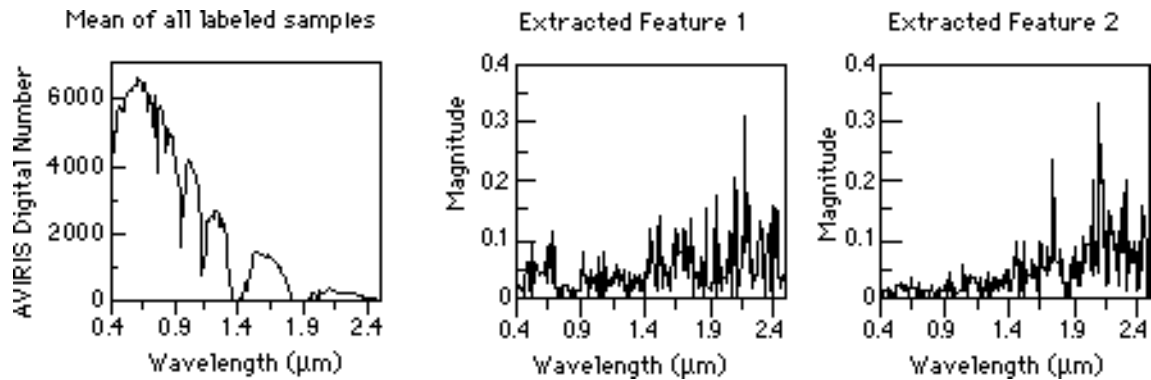


Figure 6. Average radiance and derived optimal features for discriminating between minerals of interest.

The graph on the left of Figure 6 shows the average radiance level of data gathered by the AVIRIS hyperspectral sensor. It is seen that the signal level will be quite low in the region above 2 μm , where the known diagnostic absorption features are located. The two graphs on the right of Figure 6 show the first two features extracted by the discriminant analysis feature extraction (DAFE) method. The magnitude of a DAFE feature at any specific wavelength is related to the relative significance of that wavelength for discrimination purposes. Two observations are relevant.

- (1) The DAFE features have their largest values at the narrow wavelengths where the absorption features of the minerals of interest are known to be even though the S/N in that region is low, thus confirming that these features are diagnostic. It was not necessary to specify that these features were to be used, as DAFE automatically determines which spectral regions are diagnostic and constructs a linear combination of the original bands that maximizes the separation of the specified classes.
- (2) The DAFE algorithm is able not only to respond to specific known narrow spectral features, but it can take advantage of perhaps less prominent characteristics present in the data from the entire spectrum available.

Thus, properly applied, feature extraction algorithms such as DAFE should have an advantage over such methods as the principal components transformation, which do not utilize class specific information.

On Specifying the User Classes Quantitatively

Fundamental signal processing theory dictates that, for a classifier to be well trained,

- The list of classes must be **exhaustive**, in the sense that there is a logical class to which to assign every pixel in the scene,
- The classes must be **separable** using the available features, and
- The classes must be **of informational value**, i.e. they must be classes of interest to the user.

The former of these three is required to insure that the process is a relative one rather than an absolute one. The second is a condition controlled by the data, and the third is the manner in

which the user's requirements become expressed in the analysis. The simultaneous satisfaction of these three requirements is what insures an optimal analysis and is the goal of the analyst.

An equivalent statement to these three is that a well trained classifier must have successfully modeled the distribution of the entire data set, but it must done in such a way that the different classes of interest to the user are as distinct from one another as possible. What is desired in mathematical terms is to have the density function of the entire data set modeled as a mixture of class densities, i.e.,

$$p(\mathbf{x} | \theta) = \sum_{i=1}^m \alpha_i p_i(\mathbf{x} | \theta_i)$$

where \mathbf{x} is the measured feature (vector) value, p is the probability density function describing the entire data set to be analyzed, θ symbolically represents the parameters of this probability density function, p_i is the density function of class i desired by the user with its parameters being represented by θ_i , α_i is the weighting coefficient or probability of class i , and m is the number of classes.

The problem of specifying the classes to adequate precision is perhaps the most significant aspect of the analysis process for high dimensional data. On the one hand, the large volume of high dimensional feature space, in theory, provides great potential for both the possible accuracy and the level of detail of class discrimination. However, on the other hand, it means great precision and detail is required in specifying the quantitative description of the classes desired. The expected accuracy of a given classification can be shown to be directly related to the number of training pixels used to estimate the above class density functions.

Hyperspectral Analysis Example of Urban Data

We conclude with an example analysis of an airborne hyperspectral data flightline over the Washington DC Mall. In this case the data were collected by the HYDICE sensor in 210 bands in the 0.4 to 2.4 μm region of the visible and infrared spectrum. This data set contains 1208 scan lines with 307 pixels in each scan line. It totals approximately 150 Megabytes. The primary class of interest was rooftops, however to provide an adequately exhaustive list of classes, roads, trails (graveled pathways), trees, grass, and other were added to the class list. The class other included spectral classes for water and for shadow. A simulated color infrared photograph form of image space presentation, along with the result of the analysis in thematic map form is shown on a following page as Figure 7. The challenge of the task arises principally from the fact that (a) there are many different materials used in the building roofs of the area, and they are of various ages and a variety of conditions, and (b) some of the materials used in the roofs are the same or similar to that used in the streets.

The process of analysis followed that outlined above. A software application program called MultiSpec, available to anyone at no cost from

<http://dynamo.ecn.purdue.edu/~biehl/MultiSpec/>,

was used. It contains all the necessary algorithms. In this case, the spatial resolution of the data was fine enough that training samples could be quickly labeled for a each of the classes,

including several subclasses of the class roof. Feature extraction was next used to determine a set of 10 features optimal for this task, and the classification was carried out. The result is shown in the thematic map presentation of Figure 7.

To be practical, hyperspectral data analysis must be both robust and fast. It must be usable in a wide variety of circumstances and be usable and acceptable to those of disciplines other than signal processing engineering. The analysis described was carried out on a personal desktop computer costing less than \$3000 and required less than 3 minutes of cpu time for display, feature extraction, reformatting to the optimal subspace, and classification combined. The entire process, including analyst time for labeling training samples took less than 30 minutes. No atmospheric correction or other data adjustment was used nor would it have been helpful.

The result shown is not error free, but was deemed to be good enough for most practical uses. If further reduction of errors were desirable, a second generation product could be produced by adjustment of the defined training sets, based upon a study of this initial classification.



- Groups
- background
 - Roofs
 - Road
 - Trail
 - Grass
 - Trees
 - Other

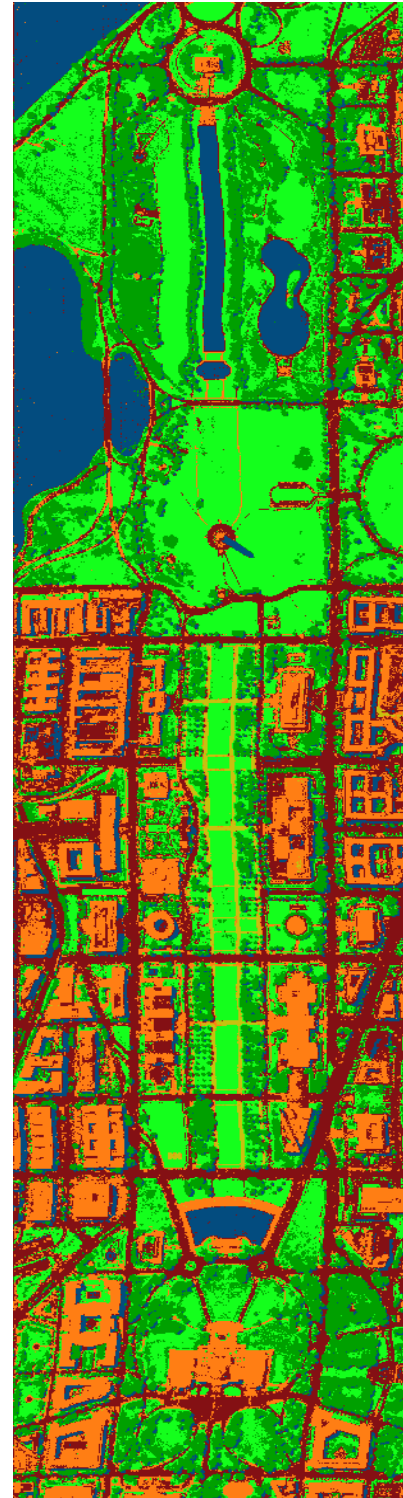


Figure 7(a) Color IR display of the data

Figure 7(b) Thematic Map result of the analysis (in color).