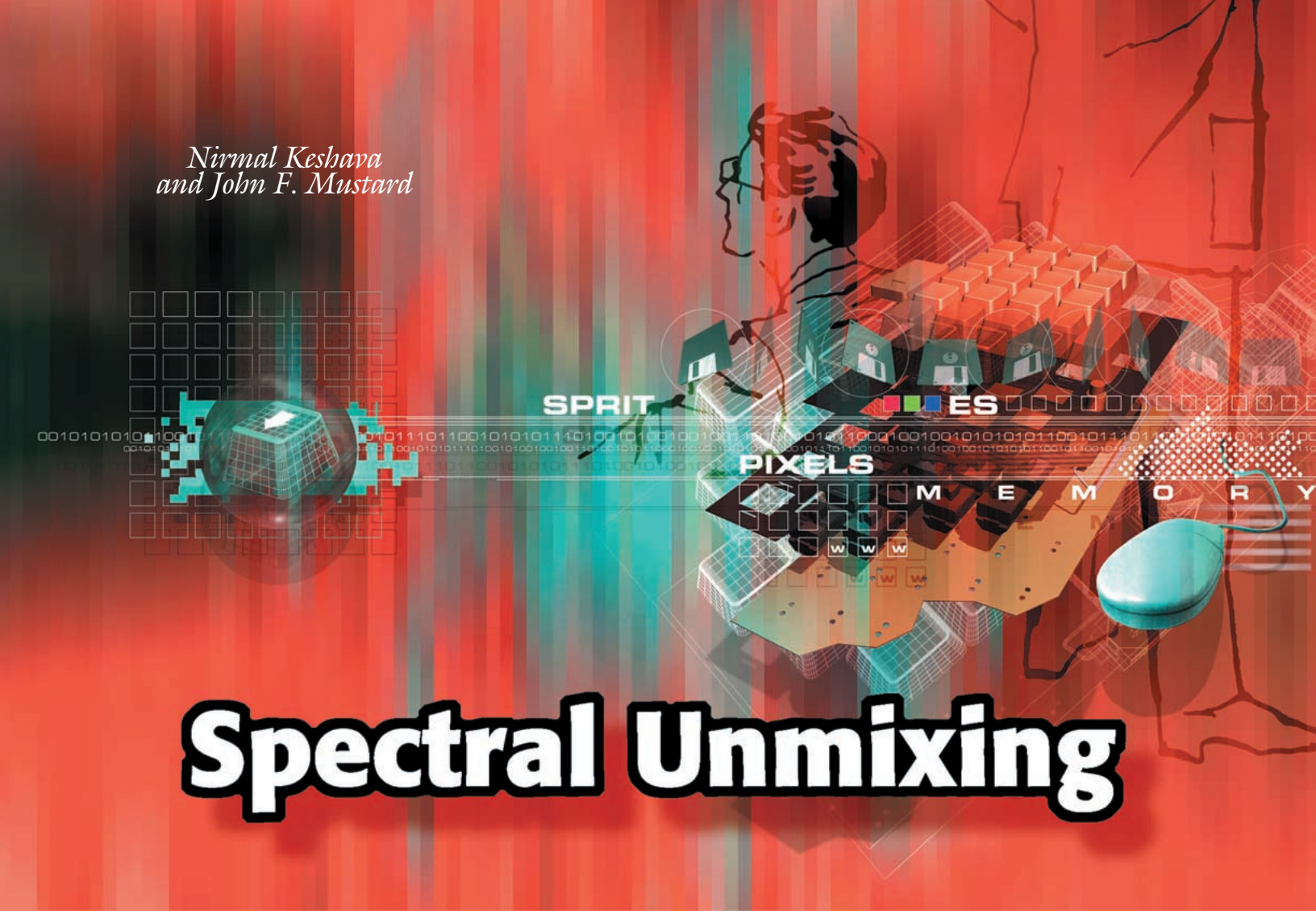
*Nirmal Keshava*
*and John F. Mustard*




# Spectral Unmixing



The evolution of passive remote sensing has witnessed the collection of measurements with significantly greater spectral breadth and resolution. It has been motivated by a desire to extract increasingly detailed information about the material properties of pixels in a scene for both civilian and military applications. While multispectral sensing has largely succeeded at classifying whole pixels, further analysis of the constituent substances that comprise a pixel is limited by a relatively low number of spectral measurements. The recognition that pixels of interest are frequently a combination of numerous disparate components has introduced a need to quantitatively decompose, or "unmix," these mixtures. Collecting data in hundreds of spectral bands, hyperspectral sensors have demonstrated the capability of performing spectral unmixing.

In hyperspectral imagery, mixed pixels are a mixture of more than one distinct substance, and they exist for one of two reasons. First, if the spatial resolution of a sensor is low enough that disparate materials can jointly occupy a single pixel, the resulting spectral measurement will be some composite of the individual spectra. This is the case for remote sensing platforms flying at a high altitude or performing wide-area surveillance, where low spatial resolution is common. Second, mixed pixels can result when distinct materials are combined into a homogeneous mixture. This circumstance can occur independent of the spatial resolution of the sensor.

*Spectral unmixing* is the procedure by which the measured spectrum of a mixed pixel is decomposed into a collection of constituent spectra, or *endmembers*, and a set of corresponding fractions, or *abundances*, that indicate the proportion of each endmember present in the pixel. Endmembers normally correspond to familiar macroscopic objects in the scene, such as water, soil, metal, vegetation, etc. Broadly speaking, unmixing is a special case of the generalized inverse problem that estimates parameters describing an object using an observation(s) of a signal that has interacted with the object before arriving at the sensor [1]. In the case of hyperspectral sensing in the reflective regime, and ignoring atmospheric effects, the incident signal is electromagnetic radiation that originates from the sun and is measured by a sensor after it has been reflected upwards by natural and man-made materials on the surface of the Earth.


1053-5888/02/$17.00©2002IEEE

Long before unmixing algorithms were investigated for remote sensing, the properties of chemical mixtures were being estimated from the diffuse and specular reflection of a sample. Reflectance spectroscopy [2] emerged in the 1920s as a complement to its sibling, transmittance spectroscopy, and was encouraged in its development by the introduction of spectrophotometers in the late 1930s and early 1940s. It soon became an established procedure for industrial inspection. Radiative transfer theory [3] was applied to analytic chemistry and reflectance spectroscopy to gauge the reflective properties of particle mixtures and the physical mechanisms that were in effect.

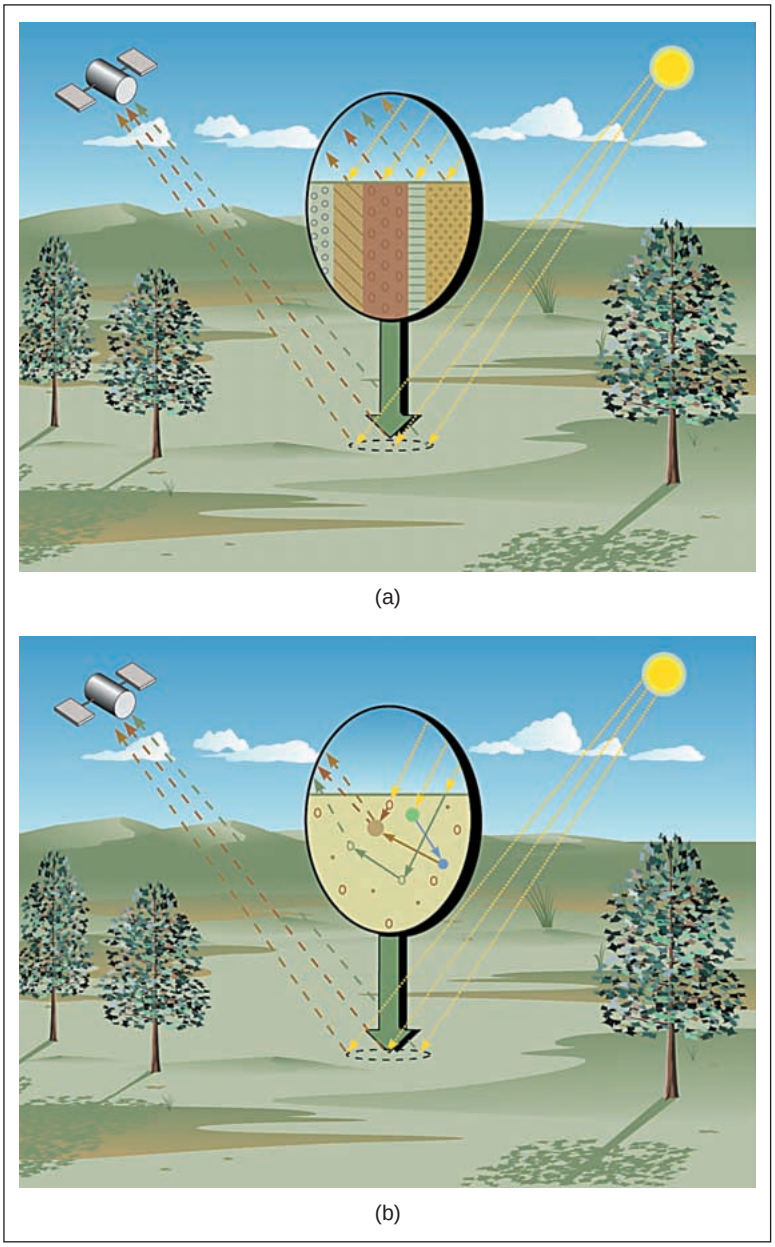
The essence of remote sensing and reflectance spectroscopy merged in the study of earth sciences using remotely sensed data for the purpose of providing synoptic analysis of geophysical phenomena [4], [5]. Together, techniques for the physical modeling of terrestrial phenomena and the subtraction of atmospheric effects permitted passive multispectral and hyperspectral radiance observations to be converted to reflectance values that described the intrinsic properties of scenes independent of the observation conditions. Thereafter, geophysicists pursued model-based methods to extract physical information from remotely sensed data by representing reflectance spectra in mathematically exploitable language [6], [7]. The result was a way to not only consistently characterize and discriminate materials on the Earth's surface, but also to decompose mixtures by spectral features.

## The Linear Mixing Model

Analytical models for the mixing of disparate materials provide the foundation for developing techniques to recover estimates of the constituent substance spectra and their proportions from mixed pixels. A complete model of the mixing process, however, is more complicated than a simple description of how surface mixtures interact. Mixing models can also incorporate the effects of the three-dimensional topology of objects in a scene, such as the height of trees, the size and density of their canopies, and the sensor observation angle.

The basic premise of mixture modeling is that within a given scene, the surface is dominated by a small number of distinct materials that have relatively constant spectral properties. These distinct substances (e.g., water, grass, mineral types) are called *endmembers*, and the fractions in which they appear in a mixed pixel are called *fractional abundances*. If most of the spectral variability within a scene is a consequence of endmembers appearing in varying proportions, it logically follows that some combination of their spectral properties can model the spectral variability observed by the remote sensing system.

In Fig. 1(a), the reflecting surface is portrayed as a *checkerboard* mixture, and any given package of incident radiation only interacts with one component (i.e., no multiple scattering between components). If the total surface area is considered to be divided proportionally according to the fractional abundances of the endmembers, then the reflected radiation will convey the characteristics of the associated media with the same proportions. In this sense, there exists a linear relationship between the fractional abundance of the substances comprising the area being imaged and the spectra in the reflected radiation.



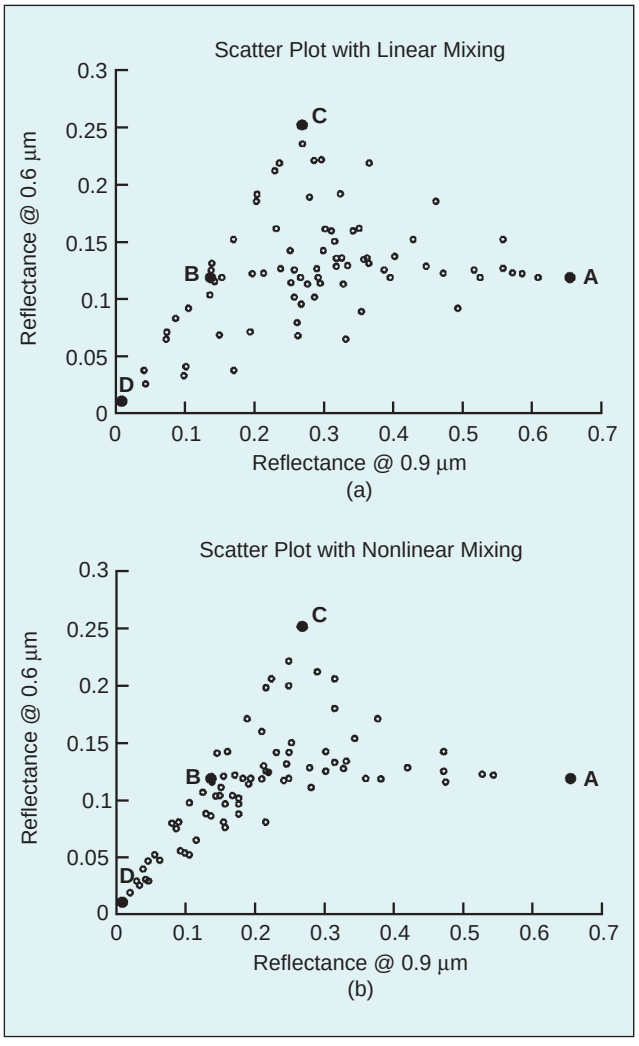

▲ *1. (a) Illustration of linear mixing where incident solar radiation reflects from surface through a single bounce and surface consists of distinct endmembers; (b) illustration of nonlinear mixing where incident solar radiation encounters an intimate mixture that induces multiple bounces.*

Hence, it is called the linear mixing model (LMM). When $M$ endmembers exist, each having $L$, and it is expressed as

$$\mathbf{x} = \sum_{\mathbf{i=1}}^{\mathbf{M}} a_i \mathbf{s}_i + \mathbf{w} = \mathbf{Sa} + \mathbf{w}$$

where $\mathbf{x}$ is the $L \times 1$ received pixel spectrum vector, $\mathbf{S}$ is the $L \times M$ matrix whose columns are the $L \times 1$ endmembers, $\mathbf{s}_i, i = 1,..., M$, $\mathbf{a}$ is the $M \times 1$ fractional abundance vector whose entries are $a_i, i = 1,\ldots, M$, and $\mathbf{w}$ is the $L \times 1$ additive observation noise vector. When $N$ pixels are considered, block notation is utilized, such that $\mathbf{X} = \mathbf{SA} + \mathbf{W}$, $\mathbf{X} = [\mathbf{x}(1) \cdots \mathbf{x}(N)]$, $\mathbf{A} = [\mathbf{a}(1) \cdots \mathbf{a}(N)]$, and $\mathbf{W} = [\mathbf{w}(1) \cdots \mathbf{w}(N)]$.

The LMM is subject to two constraints on the entries of $\mathbf{a}$. To be physically meaningful, the nonnegativity condition requires all abundances to be nonnegative such that $a_i \geq 0, i = 1,\ldots, M$. Second, as a way of accounting for the entire composition of a mixed pixel, the full additivity condition requires $\sum_{i=1}^{M} a_i = 1$.




▲ *2. (a) Linear mixing for two bands; (b) nonlinear mixing for two bands.*

### *Linear versus Nonlinear Mixing*

It has been documented that the reflectance spectrum of a mixture is a systematic combination of the component reflectance spectra in the mixture [8]. The systematics are basically linear if endmembers in a pixel appear in spatially segregated patterns, analogous to the squares on a checkerboard, as depicted in Fig. 1(a) [9]. In this case the scattering and absorption of incident electromagnetic radiation for any region on the surface is dominated by a single component on the surface, and thus the spectrum of a mixed pixel is a linear combination of the endmember spectra weighted by the fractional area coverage of each endmember in the pixel.

If, however, the components of interest are in an intimate association (the endmember materials are mixed on spatial scales smaller than the path length of photons in the mixture), as illustrated in Fig. 1(b), like sand grains of different composition in a beach deposit, light typically interacts with more than one component as it is multiply scattered, and the mixing systematics between these different components are nonlinear. Such nonlinear effects have been recognized for many years [8] in spectra of particulate mineral mixtures and are an area of active research for vegetation and canopy studies [10]. A variety of methods have been developed to treat these situations, including distinct, rigorous models for particulate surfaces [5], [11] and plant canopies [12]-[14]. This photometric model for particulate surfaces has been shown to be a powerful and useful model for nonlinear spectral mixing [5], [11]. The validity of this model for linearizing the mixture systematics has been demonstrated in laboratory studies of directional-hemispherical reflectance [15], [16] and bidirectional reflectance [17], [18] and shown to be accurate to approximately 5% absolute abundance. The technique has also been successfully applied to imaging spectrometer data for desert soils in Utah [19].

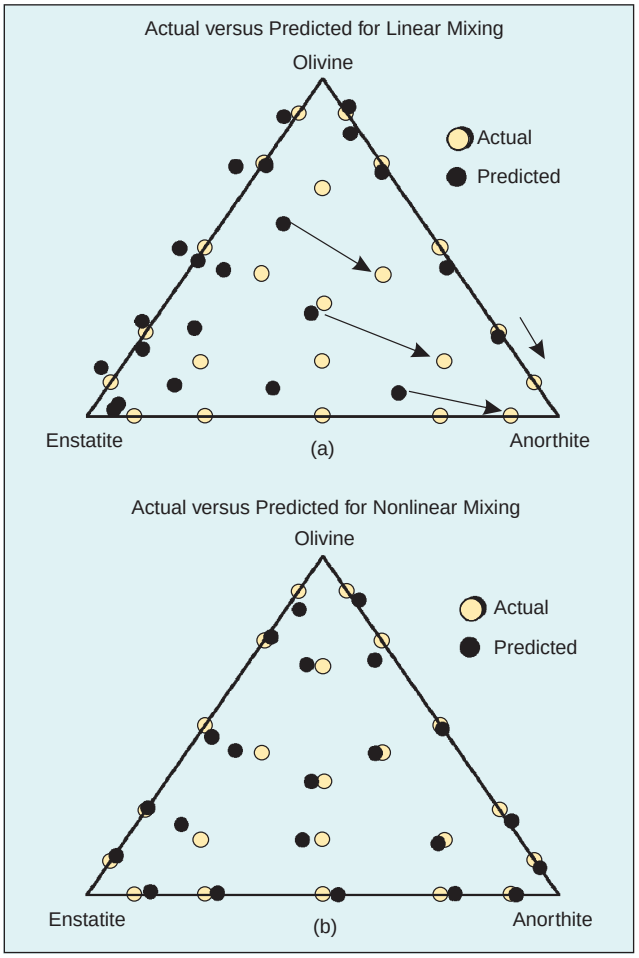
The question of whether linear or nonlinear processes dominate the spectral signatures of mixed pixels is still an unresolved issue. It likely depends on a number of factors and conditions of the scene. The linear approach has been demonstrated in numerous applications to be a useful technique for interpreting the variability in remote sensing data and a powerful means for converting spectral information into data products that can be related to the physical abundance of materials on the surface. Nevertheless, it is only strictly valid for the situation where the endmembers are arranged in discrete, segregated patches on the surface. This condition is almost never met in nature, and many constituents of interest for earth science investigations exist in soils, or at smaller scales, in intimate association with one another.

The effects of nonlinear mixing on reflectance spectra can be quite dramatic, as illustrated in Fig. 2. This is a two-dimensional plot (reflectance at 600 nm plotted against the reflectance at 900 nm) of a data cloud generated using a five-endmember mixing scenario. In Fig. 2(a), the 40 mixture points that constitute the cloud were

calculated using a linear mixing model and prescribed fractions. Along the lines joining the endmembers (e.g., line A-C), the mixtures possess 25% increments of the associated pair of endmembers. Visually, one recognizes this as a linear problem. The mixture points are spread along the lines joining the endmembers in this reflectance space at intervals inversely proportional to their fractions. Thus the point half way along a given line (e.g., line A-C) represents a mixture of 50:50 of the endmembers at the vertices of the line. The systematics are very different in Fig. 2(b) where the same five endmembers and the 40 mixture points are shown, but here the mixture spectra were calculated using the nonlinear mixing model of Mustard and Pieters [18] adapted from Hapke [11]. The nonlinear effects are clearly indicated by the curvilinear segments joining endmembers (e.g., A-C, A-D). In addition, the entire data cloud is shifted to the left against the segment C-D and towards the low albedo endmember (D). What drives this shift is the fact that during multirate scattering, low albedo materials absorb photons more readily than high albedo materials and thus lower the overall reflectance of a mixture disproportionately to their abundance in that mixture.

There are some important implications of these differences for spectral mixture analysis. If a linear mixing model is used on data where the systematics are nonlinear, the calculated fractions will be significantly in error. In tests of linear versus nonlinear mixing on laboratory data, the fractions calculated may be in error by as much as 30% absolute [40]. In addition, the linear model can cause considerable ambiguity and false fractions when used on nonlinear mixtures. Absorption bands and continua in nonlinear mixtures cannot be adequately fit with a linear model. However, the least square approach, typically used in the LMM, will minimize fitting errors using any of the endmembers in the equation. Thus, endmembers not present in a mixture will be calculated to be present simply to minimize the error. In Fig. 3, the abundances used to prepare the mixtures are shown on a ternary diagram together with calculated abundances using a linear and nonlinear mixture model. The linear model fractions in Fig. 3(a) are significantly in error, and for the enstatite-anorthite mixture, a component of olivine is predicted which is not in the actual mixtures. It is evident, however, that the nonlinear model in Fig. 3(b) accurately predicts the modal abundances of the mixtures from the reflectance spectra.

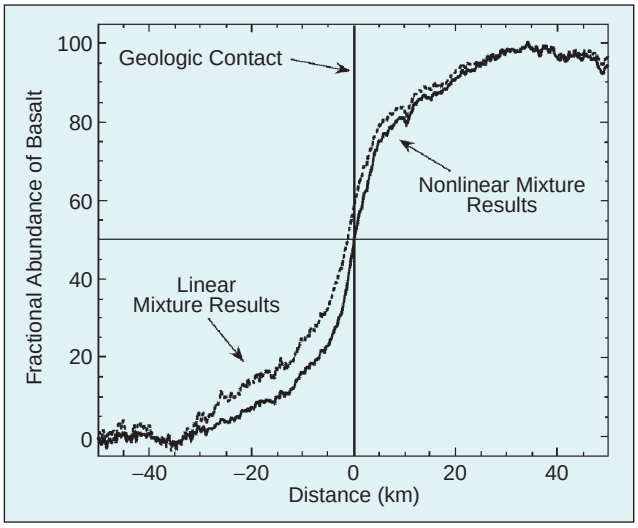
To illustrate the importance of using the proper model in an actual application, we refer to a recent study [20] that investigated geologic boundaries on the Moon that contrast ancient highland rocks adjacent to younger volcanic plains. Since the emplacement of the plains, there has been a steady bombardment of the surface by meteoroids resulting in impact craters and a redistribution of material. A central question has been to what extent this process has affected compositional boundaries. Mustard et al. [20] examined a well-preserved compositional boundary between highland rocks and the volcanic plains using images acquired by the Clementine spacecraft [21]. A multispectral camera acquired images in five spectral bands between 400 nm and 1000 nm with a spatial resolution that averaged between 120 and 150 meters/pixel. They compared results obtained from a linear mixture model with those from a nonlinear mixture model. In Fig. 4, the linear model returned a somewhat puzzling result. There was an asymmetric transport of volcanic material to the highlands (more volcanic materials to the highlands than highlands to the plains). Since the impact process is random in space and time, there is no reason to expect this type of asymmetry. The nonlinear model, however, indicated that the compositional boundary was perfectly symmetric, with equal amounts of highland to the plains as volcanics to the highlands. This symmetric distribution of material was readily modeled by diffusion equations and could be related back to the physics of the impact process. Consequently, the nonlinear results provided important constraints on the nature of this process. If the linear results had been used, however, there would have been an unfortunate need to dream up ad hoc processes to explain the result.




▲ *3. Comparison of abundance estimates for (a) linear mixing mode and (b) nonlinear mixing model.*

**Spectral unmixing is the decomposition of a mixed pixel into a collection of distinct spectra, or endmembers, and a set of fractional abundances that indicate the proportion of each endmember.**

Despite the obvious advantages of using a nonlinear approach for intimate mixtures, this has not been widely applied to remotely acquired data of particulate surfaces. There are several reasons for this. It is generally thought that the detailed scattering properties of all endmembers and surfaces are required to perform the photometric calculations. However, if we assume that most surfaces are Lambertian (i.e., scatter light equally in all directions), then the calculations are much more tractable. The Lambertian approximation is generally appropriate for particulate (e.g., soil) surfaces viewed from nadir for incidence angles up to 40°. The recovery of mixture parameters from reflectance requires knowledge of the angular orientation of the endmembers and the observation platform. Spacecraft and aircraft pointing information and digital elevation models could be used to generate this information routinely. The most important obstacle to implementing nonlinear models is that the particle size, composition, and alteration state of the endmembers are very important controlling parameters of the solutions. Nevertheless, through careful consideration of the nature of the remote data, and as spectral libraries become more well endowed with data, many of the obstacles will be surmounted.




▲ *4. Lunar compositional analysis. The nonlinear model correctly accounts for the equal distribution of material on both sides of a geologic boundary.*

## Algorithms for Linear Unmixing

Algorithms for unmixing that are discussed in the remainder of this article are focused on exploiting the LMM because it is the most frequently used model for representing the synthesis of mixed pixels from distinct endmembers. It is not difficult to see that the determination of the endmembers and the estimation of the abundances can be considered two separate problems. As a matter of fact, an examination of the literature reveals that many abundance estimators operate on the premise that the endmembers are already known, while others first seek endmembers, and yet others derive both quantities at the same time.

Whether the task is estimation of endmembers or abundances, significant attention has been focused on the computational burdens of hyperspectral processing induced by the high dimensionality of the data. As a preemptive step, some unmixing algorithms reduce the dimension of the data to sharply curtail the required computation. Not surprisingly, the familiar trade-off for less burdensome computation is decreased accuracy incurred by discarding information. We can then decompose the complete, end-to-end unmixing problem as a sequence of three consecutive procedures that are illustrated in Fig. 5 [22].

▲ *Dimension reduction:* Reduce the dimension of the data in the scene. This step is optional and is only invoked by some algorithms to reduce the computational load of subsequent steps.

▲ *Endmember determination:* Estimate the set of distinct spectra (endmembers) that constitute the mixed pixels in the scene.

▲ *Inversion:* Estimate the fractional abundances of each mixed pixel from its spectrum and the endmember spectra.

In the following sections, we examine each stage of the end-to-end unmixing problem separately and consider various approaches. We also demonstrate typical results using a coastal scene imaged by the HYDICE (HYperspectral Digital Imagery Collection Experiment) sensor in Fig. 6. The sensor collected 210 bands of data between 400 and 2500 nm for 400 lines, each having 320 samples. The data was converted from radiance measurements to reflectance values by ATREM [23]. The spatial resolution of the pixels is $1m \times 1m$. Because of several different water vapor absorption intervals, some bands have been discarded, and, consequently, 144 bands have been retained for processing. Subsequent plots of spectra retain gaps in the position of band wavelengths where values have been discarded due to poor SNR in the intervals having water vapor absorption.

## Dimension Reduction

As mentioned earlier, dimension reduction, by itself, is not a necessary step for unmixing, but because these techniques often expedite subsequent processing, it is natural to retain dimension reduction as a stage of unmixing. We

have intentionally differentiated between dimension reduction algorithms and compression algorithms to highlight an important distinction. Dimension reduction algorithms do not reduce the dimension of data with the goal of reconstructing an approximation to the original signal. Instead, they seek a minimal representation of the signal that sufficiently retains the requisite information for successful unmixing in the lower dimension. Dimension reduction algorithms are designed to minimize errors in the procedures performed in the lower dimension.

### *Principal Component Analysis*

Principal component analysis (PCA) identifies orthogonal axes for dimension reduction by performing an eigendecomposition of the sample covariance matrix of the data

$$\hat{\Gamma}_x == \frac{1}{N}\sum_{n=1}^{N}\left(\mathbf{x}(n)-\hat{\mu}_x\right)\left(\mathbf{x}(n)-\hat{\mu}_x\right)^T$$

where $\hat{\mu}_x$ is the mean vector of the pixel set. The resulting eigendecomposition can be expressed as $\hat{\Gamma}_x = \mathbf{U}\Sigma U^T$, where $\mathbf{U}$ is a unitary matrix of eigenvectors and $\Sigma$ is a diagonal matrix of eigenvalues.

The magnitude of the eigenvalues indicates the power residing in the data along the component of the data parallel to the associated eigenvector. Hence, the effective dimensionality of the data can be estimated by counting the number of nonzero eigenvalues. Premultiplication by a subset of the rows in $\mathbf{U}^T$ moves $\mathbf{x}$ to a new system of decorrelated variables oriented along the eigenvectors in $\mathbf{U}$ and results in a lower-dimensional multivariate random vector that still conveys most of the energy in the original, higher-dimensional system.

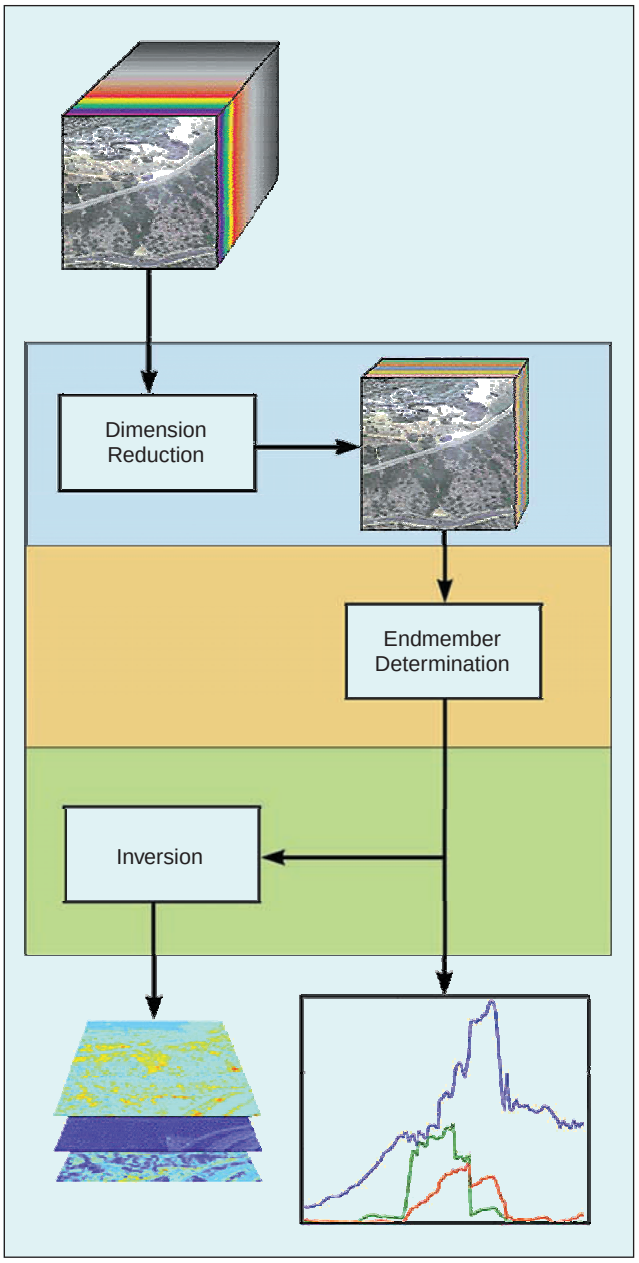
As an example of the properties belonging to a typical scene of hyperspectral data, we use data from the coastal scene in Fig. 6. The cumulative normalized eigenvalues for the scene are plotted in Fig. 7, indicating that four eigenvalues are sufficient to account for over 99% of the total energy in the scene. It is worth noting that PCA possesses optimal properties for retaining the energy in random signals, but in no way guarantees that the resulting transformation will preserve the information that aids in detecting low probability objects.

#### *Maximum Noise Fraction/Noise Adjusted PCA*

PCA works independently of any estimates of the noise, statistical or otherwise, in the signal model and, consequently, does not construct its eigenvectors in ways that optimize a signal-to-noise criterion. Two options for dimension reduction that are known by different names, but are mathematically equivalent, include statistical models for the noise in their construction of a signal transform. Maximum noise fraction (MNF) [24] or noise adjusted principle components (NAPC) [25] requires covariance information about the sources of additive noise in addition to the covariance of the data. If the estimated noise covariance is given by $\hat{\Gamma}_w$ and the estimated received signal covariance is $\hat{\Gamma}_x$, the component of $\mathbf{x}$ that possesses the maximum fraction of noise, $\mathbf{v}$, maximizes

$$\frac{\mathbf{v}^T\hat{\Gamma}_w\mathbf{v}}{\mathbf{v}^T\hat{\Gamma}_x\mathbf{v}}.$$

The axes that define $\mathbf{v}$ are the left-hand eigenvectors of $\hat{\Gamma}_w\hat{\Gamma}_x^1$, and, unlike the axes for PCA, are not necessarily orthogonal. They do, however, identify and order the components of the received signal possessing the maximum SNR.




▲ *5. Conceptual diagram of end-to-end spectral unmixing.*

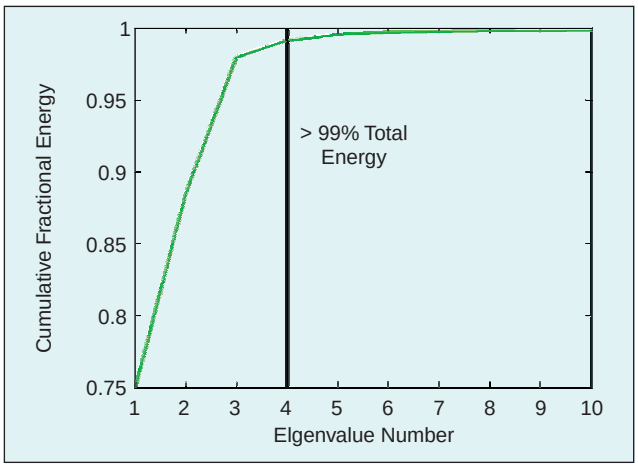
### Nonstatistical Methods

The computational complexity associated with applying statistical techniques is often considered an obstacle to real-time implementations. In contrast, algorithms for dimension reduction have been developed that avoid algorithmically expensive operations and are therefore suitable for simple and efficient hardware implementations.

#### Exemplar Selection

A nonstatistical technique for reducing the dimension of hyperspectral data appears as part of the Naval Research Laboratory's ORASIS [26] (Optical Real-time Adaptive Spectral Identification System), which is a series of hyperspectral processing modules. Before submitting the data to modules that estimate endmembers, the dimension of the hyperspectral data is reduced significantly by the exemplar selector module (ESM). When a new pixel is collected from the scene by the sensor, its spectrum is compared to pixels in the existing set of exemplars. If its distance—the usual measure of distance is the angle between two spectra (see "Metrics for Hyperspectral Data")—from each of the exemplars (the first pixel in a scene automatically becomes the first exemplar) exceeds a prescreen threshold, it is added to the collection. If it is not sufficiently different, the exemplar set remains unchanged. At periodic intervals, the salient selection module (SSM) then orthogonalizes the current set of exemplars using a modified Gram-Schmidt process and, based on an acceptable representation error, retains a subset of the orthogonal vectors to transform the data to a lower dimension.

▲ 6. Coastal scene imaged by HYDICE sensor.




▲ 7. Cumulative fractional eigenvalue energy.

## Endmember Determination

In this section, we address one of the two primary tasks of unmixing, endmember determination. The objective of these algorithms is to determine the constituent spectra that occupy the columns of **S** in the LMM. We consider this important procedure from two vantage points. The first is the empirical estimation of endmembers from a scene. This approach is concerned with the characteristics of endmembers encountered when they are estimated empirically through observation and physical intuition. It provides a context for what automated endmember estimation should strive to achieve, namely that endmembers should first possess the basic properties of realistic spectra. Second, we discuss the various approaches developed for automated endmember determination techniques. In contrast to empirical methods, automated algorithms may employ statistics to capture variability, but their analytical determination may result in endmember estimates that satisfy some optimality criterion, but are physically unrealistic.

### Interactive Endmember Determination

The key to linear unmixing is to define a set of spectral endmembers that are representative of physical components on the surface and that encompass the spectral variability inherent in a given scene. To ensure the uniqueness of a solution to the LMM, the set of endmembers must be linearly independent. Along with the additivity constraint imposed on **a**, each spectral band provides a linear constraint, such that the theoretical limit to the number of endmembers in a scene is equal to the number of spectral channels plus one. However, the number of endmembers that may be practically identified and used is far fewer, typically ranging from three to seven, depending on the number of channels and the spectral variability of the scene components. Channel-to-channel variance in spectral data sets is highly correlated, thus some of the spectral

information is redundant. In addition, materials of different physical composition may exhibit similar spectral properties over a given wavelength range or have spectral properties that can be mathematically defined by linear combinations of other components. Such materials would therefore not be resolved. Strategies to overcome these limitations, such as multiple endmember models, have been discussed [27].

A practical way to approach the definition of endmembers is to apply a field-based framework. Given a scene to unmix, one might recognize classes of materials such as soil, vegetation, rock types, etc., and begin by selecting one endmember from each class (provided that the spectral range and resolution are sufficient to resolve them) and subsequently build complexity into the model as is warranted by further analyses. This is illustrated conceptually in Fig. 8, where we consider a scene sampled at the spectral and spatial resolution of the Landsat Thematic Mapper that consists of three basic components: vegetation, a rock unit represented by gabbro, and shade. The reflectance values for six bands are plotted in Fig. 8(a). The shade endmember is one of the more interesting and novel features of spectral mixture analysis. Natural surfaces are never uniformly illuminated, and variations are caused by topographic slope changes which result in both shading and shadow that vary with the seasons and the diurnal cycle, as well as shadows at subpixel scales caused by small-scale topography, trees, shrubs, and boulders. The shade endmember mixes with the other endmembers in the scene in proportion to the amount of this variation in illumination and may be considered as a neutral multiplicative scaling factor.

The first stage in a mixture analysis is to define a suite of image endmembers (selected from the image data). An image endmember (IE) is obtained by locating a pixel in the scene with the maximum abundance of the physical endmember it will represent. This is illustrated in Fig. 8(b) where the IE are represented by the open circles. They may be selected using algorithms [27], [31], [32] or based on criteria such as field knowledge or other analysis methods (e.g., ratios, PCA). For a typical scene, it is unlikely that a single pixel or group of pixels can be selected

## Metrics for Hyperspectral Processing

A quick survey of the algorithms currently used for spectral unmixing reveals a mix of approaches that borrow liberally from well-known traditional approaches but also integrate newer, more physical interpretations. While some techniques have demonstrated significant success, it is still unclear exactly what combination of physical and mathematical modeling can optimally extract information from hyperspectral signals.

At the core of algorithms are distance metrics that define the notion of similarity between two signals. The comparison of two hyperspectral signals can be performed in many ways, but two complementary measures based on the Euclidean formulation have been commonly employed. For two spectra, $\mathbf{x}$ and $\mathbf{y}$, the spectral angle mapper (SAM) measures the angle, $\theta$, between them

$$\theta(\mathbf{x},\mathbf{y}) = \arccos\left(\frac{<\mathbf{x},\mathbf{y}>}{\|\mathbf{x}\|_2\|\mathbf{y}\|_2}\right) \quad 0 \le \theta \le \frac{\pi}{2}.$$

The Euclidean minimum distance (EMD) is the two-norm of $(\mathbf{x}-\mathbf{y})$ and calculates the length, $\Delta$, of the shortest path between two vectors

$$\Delta(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|_2.$$

Both metrics have found applicability for different reasons. SAM is invariant to the multiplication of $\mathbf{x}$ and $\mathbf{y}$ by constants and, consequently, is invariant to unknown multiplicative scalings of spectra that may arise due to differences in illumination and angular orientation. Likewise, EMD is the foundation for least squares analysis, which has a strong historical legacy and is frequently used for unmixing.

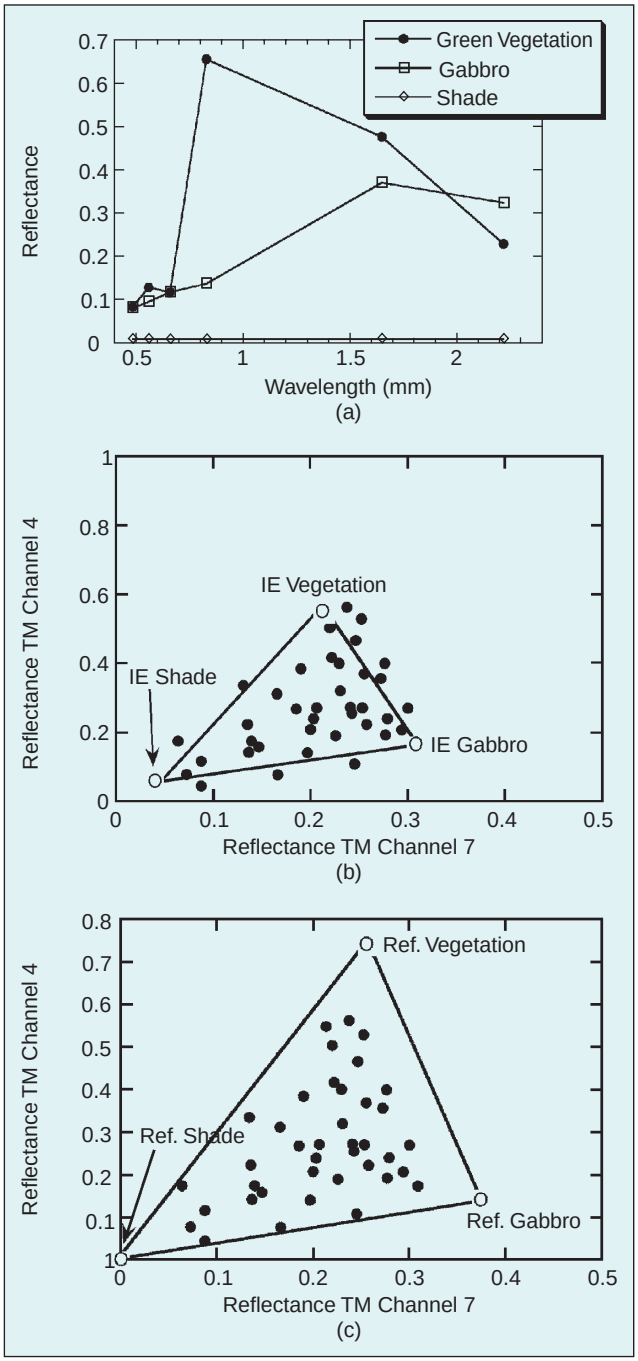
No single metric is useful in all circumstances, but the fundamental issue confronting the selection of a distance metric is determining a physically meaningful way of quantifying the similarity between two hyperspectral signals. From a physical viewpoint, SAM and EMD give no indication of similarities in physical or material properties, instead providing a single scalar metric to characterize the similarity of two vectors. Yet, they provide the foundation for most algorithms in the hyperspectral community.

Hyperspectral data possesses two notable properties that should be taken into account when contemplating metrics. First, hyperspectral signals are vectors with nonnegative entries. Second, although a pixel spectrum may be represented in vector form, its entries are not equal from a radiometric viewpoint. In fact, widths of bands are nonuniform and, unless the data has been appropriately normalized, equal values from bands having unequal widths indicate different radiometric quantities. Hyperspectral measurements, in short, are radiance measurements organized in a vector, but there is much more information to be exploited than just the vector entries alone.

The obvious question becomes "What should the notion of distance in hyperspectral processing convey?" One answer, as confusing as it sounds, is that it should convey whatever the application demands of it. The crux of hyperspectral processing is in mapping mathematical structures and operators to the physical processes that give rise to the radiance arriving at the sensor through the intervening atmosphere. The debate between linear and nonlinear mixing models is indicative of the fact that the technique for unmixing—and, hence, the choice of metric—should reflect the physics of the mixing process. It may be impossible to design the perfect metric for every physical circumstance, but the opportunity, and motivation, to close the loop between the mathematical and physical world certainly exists in hyperspectral processing.

that will provide a pure component (the pixel is composed of 100% of the endmember), though there will likely exist some pixels that are nearly pure. Conceptually, the image endmembers are those that best bound the data cloud as shown in Fig. 8(b).

The fractions calculated for each of the data points (filled circles) in Fig. 8 correspond to the relative distance from the IE points, subject to the full additivity constraint. Since the image endmembers are themselves most likely mixtures of other materials, there may exist fractions in the results that will be greater than one or less than zero, although the sum of the fractions will still equal one. Data points that fall within the lines joining the IEs have positive fractions, while those that fall outside the lines have negative fractions of the IE opposite this line. For example, the pixels that plot outside (to the right) of the line joining the image endmembers gabbro and vegetation will have negative shade fractions. Thus, negative fractions, or fractions that are greater than 1.0 do not necessarily indicate an error in the method of the application.




▲ *8. Endmember analysis of six-band spectra. (a) Reflectance spectra for two substances and shade; (b) image endmembers that tightly bound the variability but cause some spectra to reside the range of acceptable abundance values; (c) reference endmembers that enclose all spectra in a scene.*

The results using image endmembers provide a first-order perspective on the mixing relationships, and if the endmembers are relatively pure, then the estimated abundances will be similar to the actual abundances. However, it is often desirable to use library or reference endmembers to provide a better link to spectral libraries and therefore ground-truth. Moreover, using a standard set of reference endmembers is useful when abundance estimates are to be used in multitemporal studies or across scenes that were acquired under different illumination conditions or times of years. This is illustrated in Fig. 8(c) where it can be seen that the reference endmembers now fully bound the variability of the image data. In this example, there would be no negative abundances, or abundances greater than 1.0.

## Automated Endmember Determination

In interactive endmember determination, image endmember selection is achieved through an educated trial-and-error approach. An analyst has some knowledge of the field site or data set, and a set of objectives for conducting the analysis. In many situations, however, an automated method of determining these essential components is desired. For example, results should be repeatable, and the fraction images should describe realistic physical variables or components in the scene. Several techniques have been developed to estimate endmembers that do not require specific assumptions on the probabilistic densities of the data.

### *Nonparametric Methods*

Clustering algorithms have been adapted to infer estimates of $\mathbf{S}$ by identifying natural partitions in data that exemplify distinct statistical behavior [28]. They can be useful when the only information available is a set of mixed pixel spectra, $\mathbf{X}$, from a scene. Several variations on traditional clustering algorithms, such as K-means clustering, have attempted to incorporate the LMM into the problem formulation, such that the resulting centroids serve as estimates of endmembers. Similarly, as an extension of classification algorithms that assign the class label of the nearest centroid to pixels, abundance estimates that implicitly observe the nonnegativity and full additivity conditions are derived from the relative proximity of a pixel to each centroid [29].

### *Parametric Methods*

Parametric algorithms have added the additional assumption of Gaussianity to the endmember determination problem, but at the same time have attempted to incorporate a strong physical intuition into purely statistical methods. The characteristics of linear mixing have been extended to a *stochastic mixing model* (SMM) where mixed pixels in a scene are modeled as combinations of fundamental, or hard, endmembers. The fundamental endmembers are identified from the data and modeled as Gaussian distributions. These techniques resemble the analysis of Gaussian mixtures, with maximum likelihood (ML) estimates of the class parameters corresponding to each hard endmember derived iteratively using the expectation maximization (EM) algorithm [30].

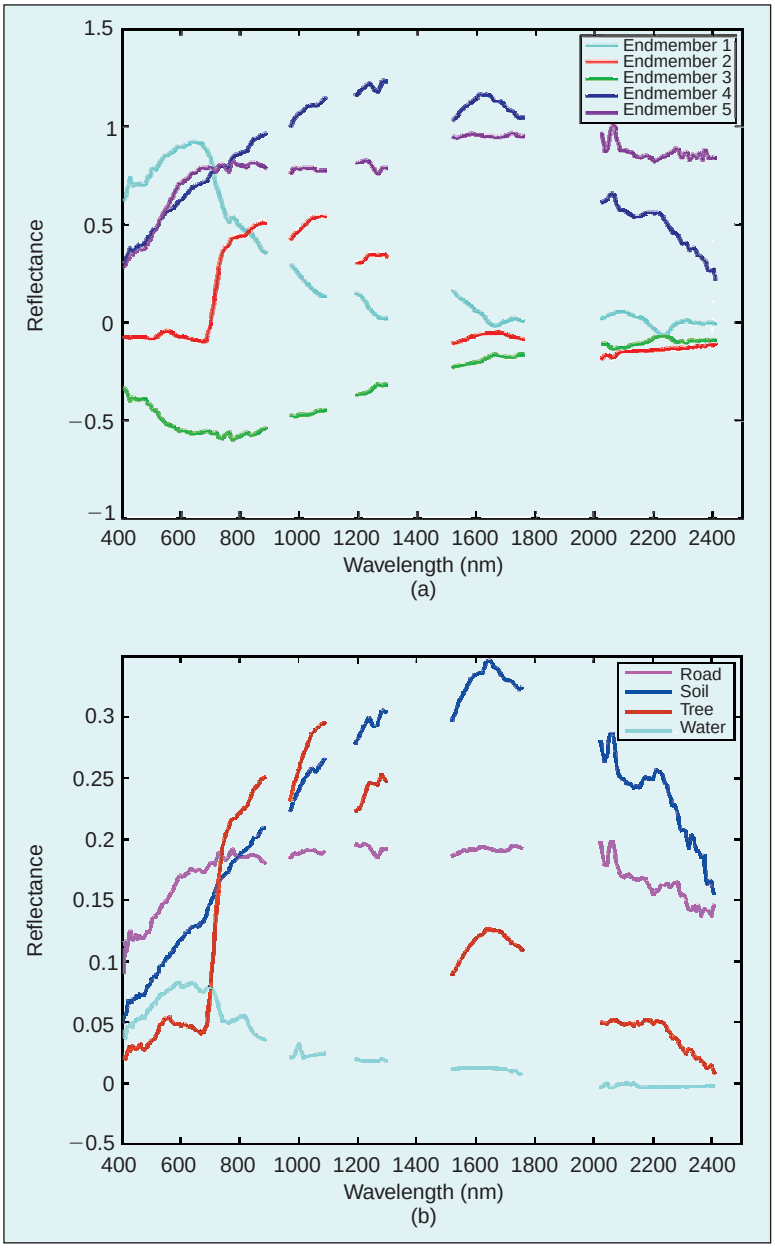
Another parametric approach attempts to incorporate a priori physical information into a statistical framework. In modified spectral mixture analysis (MSMA) [31], the endmembers and abundances are assumed to be deterministic and unknown. The simultaneous recovery of ML estimates for both **S** and **a** results in a nonlinear formulation that employs an iterative, damped least squares technique. Most importantly, the iterations begin with initial guesses for **S** and **a**, where a priori knowledge about endmembers and abundances may be inserted, and their confidence can be qualified by an accompanying covariance. Successive estimates of MSMA minimize the error in the fit and constraints on the values of abundances and endmembers are enforced between iterations.

### *Geometric Endmember Determination*

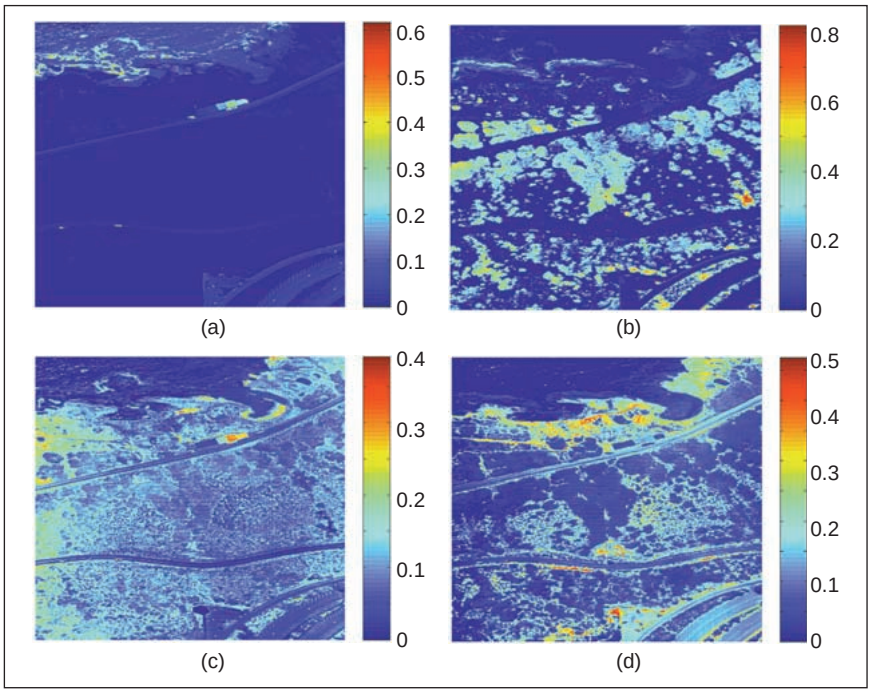
Interactive methods rely primarily upon physical intuition and human supervision to derive endmembers, whereas statistical methods are repeatable, statistically optimal, and automated. Both approaches have advantages and disadvantages. In contrast, geometric approaches [32], [33] exploit the strong parallelism between the LMM and the theory of convex sets. The fundamental assumption of geometric endmember determination is that endmembers are pure substances in a scene whose spectra reside at the extremities of the volume occupied by the data, and, consequently, all mixed pixels that occupy the interior of the volume are linear mixtures of endmembers.

For a given scene, if there is sufficient mixing of endmembers to populate the extreme planes of the volume, then estimates of endmember spectra may be derived from the vertices of the multifaceted simplex that most tightly encloses the data and has the same number of endmembers as vertices. The procedure that performs this optimization is known as *shrinkwrapping* [32]. Because hyperspectral data has a high number of dimensions, shrinkwrapping is never performed on the original spectra, but is instead performed on a lower-dimensional representation. Techniques such as MNF and PCA are often utilized to identify the axes for this dimension reduction. Moreover, since shrinkwrapping only requires knowledge of the pixels on the perimeter of the volume, a convex hull operation is performed on the data after dimension reduction and before shrinkwrapping.

Fig. 9(a) plots the significant endmember reflectance spectra derived from the scene in Fig. 6 using a geometric analysis that employs shrinkwrapping. For comparison,




▲ *9. (a) Endmember reflectance spectra derived using geometric endmember determination; (b) reflectance spectra of different materials appearing in scene.*




▲ *10. Fraction planes for (a) Endmember 1, (b) Endmember 2, (c) Endmember 4, and (d) Endmember 5 derived using the nonnegative least squares algorithm.*

sample spectra from visibly distinct materials in the same scene are plotted in Fig. 9(b). Endmembers 1, 2, 4, and 5 bear a strong resemblance in their overall shape to the water, tree, soil, and road spectra in Fig. 9(b). Endmember 3, by virtue of being mostly negative, is a noise endmember.

The magnitudes of Endmembers 1, 2, 4, and 5 differ significantly from the spectra in Fig. 9(b) for two reasons. First, there is a considerable amount of variability in the spectra from the same object. Adjacent pixels made from the same material may be illuminated differently due to shadowing, or angular orientation and, as a consequence, the reflectance spectra possess the same shape, but are scaled differently. Second, the shrinkwrapping algorithm must enclose all pixels, meaning that its vertices must be positioned such that the simplex captures every pixel in the scene. Endmember variability or outliers may force a vertex to reside at a point beyond the nominal position of the endmember in order to enclose every point. This effect is evident in Fig. 9(a) where the individual reflectance values sometimes exceed one or become negative, which is physically impossible.

While the geometric model for linear unmixing does seem appropriate, it also has limitations. In contrast to the statistical dimension reduction algorithms that capture general trends in data, but can also overlook important, low- probability events, shrinkwrapping is susceptible to outliers and artifacts that may adversely change the shape of the simplex and hence, the estimate of the endmembers. Reconciling the results of geometric techniques with statistical approaches may remedy or mitigate this problem.

## Inversion

In the previous section, we discussed endmember determination and discovered that it is often interrelated with estimating the abundance vector, **a**, in the LMM. A key aspect of inversion is the incorporation of the dual physical constraints that **a** must obey, full additivity and nonnegativity. Quite often, the basis for arriving at estimates is a distance metric that is minimized. As previously discussed, methods for estimating endmembers based on the SVD, as well as many of the statistical methods minimize some quantity related to squared-error. The same characteristic holds true for a significant majority of inversion algorithms.

### *Least Squares Methods*

The class of inversion algorithms based on minimizing squared-error start from the simplest form of least squares inversion and increase in complexity as further assumptions and parametric structure are imposed on the problem. Variations of the least squares concept have been adopted to reflect the unique circumstances associated with hyperspectral data.

#### *Unconstrained Least Squares*

Starting with the LMM, and the assumption of no additive noise, the unconstrained least squares solution for $a$ is $\hat{\mathbf{a}}^U = (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{x}$ [34]. This unconstrained estimate for **a** minimizes $|\mathbf{x} - \mathbf{S}\hat{a}^U|^2$. This form requires no estimate of the additive noise and exists when there are more bands than endmembers (a reasonable assumption for hyperspectral sensing), and when **S** has full column rank.

#### *Full Additivity*

The unconstrained solution can be refined by constraining $\hat{\mathbf{a}}$ to fulfill the full additivity condition. This has the effect of restricting the least squares solution to lie on the hyperplane given by $\sum_{i=1}^{M} a_i = 1$.

The solution for a general least squares estimate having linear constraints given by $\mathbf{Za} = b$, is obtained using Lagrange multipliers and is given by [35]

$$\hat{\mathbf{a}}^F = \hat{\mathbf{a}}^U - (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{Z}^T\left[\mathbf{Z}\left(\mathbf{S}^T\mathbf{S}\right)^{-1}\mathbf{Z}^T\right]^{-1}\left(\mathbf{Z}\hat{a}^U - \mathbf{b}\right).$$

The full additivity constraint is enforced when **Z** is a $1 \times M$ row vector having all ones and $\mathbf{b} = 1$. Closer examination of $\hat{\mathbf{a}}^F$ reveals that the solution enforcing full additivity consists of the unconstrained least squares solution, $\hat{\mathbf{a}}^U$, with an additive correction term that depends

on the matrix of endmembers, $\mathbf{S}$, and the error incurred by $\hat{\mathbf{a}}^U$ in satisfying the full additivity constraint.

*Nonnegativity*

The complementary constraint, nonnegativity, is not as easy to address in closed form as full additivity. Minimizing $|\mathbf{x}-\mathbf{S}\mathbf{a}|^2$ while maintaining $a_i \geq 0, i=1,\ldots,M$, falls in the domain of quadratic programming with linear inequalities as constraints. An alternative that has been employed in practice is the nonnegative least squares (NNLS) algorithm [36]. The approach here is to iteratively estimate $\mathbf{a}$, and at every iteration, find a least squares solution for just those coefficients of $\mathbf{a}$ that are negative using only the associated columns of $\mathbf{S}$. In contrast to $\hat{\mathbf{a}}^U$, $\hat{\mathbf{a}}^{\mathrm{NN}}$ is a constrained solution that minimizes, $|\mathbf{x}-\mathbf{S}_{\mathrm{POS}}\mathbf{a}|^2$, where $\mathbf{S}_{\mathrm{POS}}=\mathbf{S}$ for those columns of $\mathbf{S}_{\mathrm{POS}}$ where the associated entry of $\hat{\mathbf{a}}^{\mathrm{NN}} > \mathbf{0}$, and the zero vector appears in columns where $\hat{\mathbf{a}}^{\mathrm{NN}} = \mathbf{0}$.

In Fig. 10, four fraction abundance planes depict the fractional abundance values derived using the NNLS algorithm for Endmembers 1, 2, 4, and 5. The fraction planes exhibit large values where the corresponding endmember occupies large fractions of pixels and small values elsewhere. As expected, Endmember 1 exposes the mixed nature of the ocean water, where regions closer to shore are illuminated more prominently than waters that are farther from shore. Part of this can be accounted for by the fact that pixel spectra from the ocean water show a large degree of variability in magnitude, while retaining the same basic shape (see Endmember 1 in Fig. 9(a)). Also, the additional reflectance from the ocean floor contributes to the greater magnitudes in shallow water.

Fig. 11 shows, however, that despite constraining the abundance values to be nonnegative, they rarely sum to one. The optimization designed to assign abundances for each endmember to every pixel in the scene succeeds at obtaining purely nonnegative values, however, it is unable to assure that they always sum to one. Generally, the results in Fig. 10 demonstrate that simple inversion algorithms based on familiar concepts in signal processing are capable of exposing similarities to multiple endmembers, but more sophistication and physical reasoning is necessary to arrive at confident estimates on a pixel-by-pixel basis.

## Minimum Variance Methods

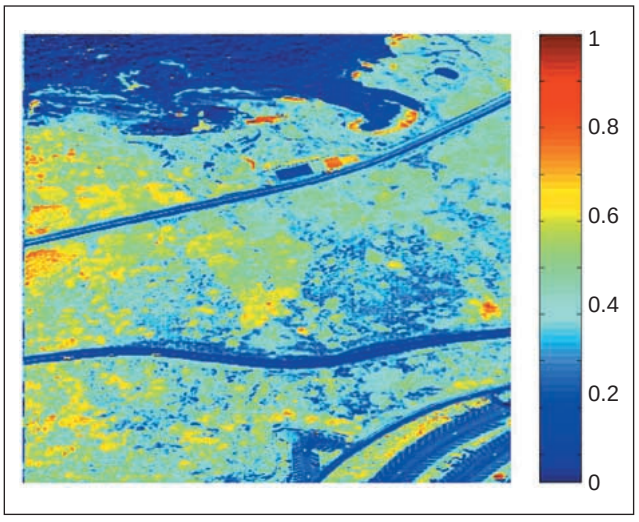
In the previous examples of least squares inversion algorithms, the common objective has been to estimate abundances that minimize the squared-error between the actual spectrum and the approximated spectrum. The statistical analogue of least squares estimation minimizes the variance of the estimator. Under the assumption that the additive noise vector, $\mathbf{w}$, in the LMM is a zero-mean random process and has a covariance, $\Gamma_w$, the minimum variance estimate of the abundances, $\hat{\mathbf{a}}^V$, is $\hat{\mathbf{a}}^V = (\mathbf{S}^T\Gamma_w^{-1}\mathbf{S})^{-1}\mathbf{S}^T\Gamma_w^{-1}\mathbf{x}$. Because this unbiased estimator

**The LMM is the most frequently used model for representing the synthesis of mixed pixels from distinct endmembers.**

is the best linear unbiased estimator (BLUE), it is called the minimum variance unbiased estimator (MVUE). Further, if $\mathbf{w}$ is Gaussian, in addition to the aforementioned properties, it achieves the Cramer-Rao lower bound for unbiased estimators [37].

## Variable Endmember Methods

Finally, it is important to recognize a class of algorithms that inverts individual pixels using different sets of endmembers. A central precept of these techniques is that the number of endmembers required to unmix an entire scene can be considerably greater than the number present in any individual pixel, and each pixel in a scene may utilize a different subset of endmembers. To cull a library of endmembers, a criterion for model fitness must work to restrict the number of endmembers employed to invert an individual pixel. One method performs unconstrained least squares inversion, discards those endmembers yielding negative values, and then performs the inversion again minus the discarded endmember, repeating the entire process until all endmembers still present have nonnegative abundance values. To enforce the additivity condition, the sum of abundances is then normalized to one [38]. A second method works in the opposite direction by first finding the best pair of endmembers and then incrementally adding additional endmembers so long as certain measures of model fitness are met [39].




▲ *11. Sum of abundances using NNLS algorithm.*

## Discussion

While many techniques for unmixing have been introduced and have shown promise in experiments, precise verification of algorithm results with real data has been elusive. The primary reason has been the lack of ground truth. Unlike other applications, accurately gauging the performance of unmixing algorithms requires a detailed knowledge of the components comprising each pixel in a scene. For this reason, endmember determination has been the primary focus of experiments with hyperspectral scene data, rather than inversion. Further experiments will surely continue to investigate unmixing in greater depth.

## Summary

Spectral unmixing using hyperspectral data represents a significant step in the evolution of remote decompositional analysis that began with multispectral sensing. It is a consequence of collecting data in greater and greater quantities and the desire to extract more detailed information about the material composition of surfaces. Linear mixing is the key assumption that has permitted well-known algorithms to be adapted to the unmixing problem. In fact, the resemblance of the linear mixing model to system models in other areas has permitted a significant legacy of algorithms from a wide range of applications to be adapted to unmixing. However, it is still unclear whether the assumption of linearity is sufficient to model the mixing process in every application of interest. It is clear, however, that the applicability of models and techniques is highly dependent on the variety of circumstances and factors that give rise to mixed pixels.

The outputs of spectral unmixing, endmember, and abundance estimates are important for identifying the material composition of mixtures. Unmixing is a close relative to another important problem in hyperspectral processing, the subpixel target detection problem. Mixtures, however, are not limited to simple terrestrial components. Unmixing has also been applied to the analysis of mixed gases in the longwave infrared (5,000-14,000 nm). Ever-growing spectral libraries that reduce the dependence on in situ determinations of endmembers have aided these efforts. Nevertheless, the pursuit of increasingly accurate and precise knowledge from remote sensors puts spectral unmixing on the forefront of future remote sensing missions and research endeavors.

## Acknowledgments

We would like to thank MIT Lincoln Laboratory staff members Michael K. Griffin for his expertise on atmospheric compensation, Craig Richard for performing the atmospheric compensation, and Peter W. Boettcher for implementing the geometric endmember determination algorithms. We would also like to thank Jeffrey Bowles of the Naval Research Laboratory for providing an update on the description of the ORASIS algorithm as well as Stefanie Tompkins of SAIC for useful comments.

*Nirmal Keshava* has been a technical staff member at Lincoln Laboratory since 1997. He received the B.S. from UCLA, the M.S. from Purdue University, and the Ph.D. from Carnegie Mellon University, all in electrical engineering. His current research interests focus on physical and mathematical modeling for the exploitation of hyperspectral imagery.

*John F. Mustard* is an Associate Professor of geological sciences at Brown University. He received a B.Sc. from the University of British Columbia and a Ph.D. from Brown University. He has worked for over 15 years in the application of remote sensing to Earth and planetary science combining theory, laboratory work, and field investigations.

## References

[1] A. Tarantola and B. Valette, "Generalized nonlinear inverse problems solved using the least squares criterion," *Rev. Geophysi. Space Phys.*, vol. 20, pp. 219-232, 1982.

[2] W.M. Wendlandt and H.G. Hecht, *Reflectance Spectroscopy*. New York: Interscience, 1966.

[3] S. Chandrasekhar, *Radiative Transfer*. New York: Dover, 1960.

[4] R.N. Clark and T.L. Roush, "Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications," *J. Geophys. Res.*, vol. 89, no. B7, pp. 6329-6340, 1984.

[5] B. Hapke, *Theory of Reflectance and Emittance Spectroscopy*. Cambridge, U.K.: Cambridge Univ. Press, 1993.

[6] F. Tsai and W. Philpot, "Derivative analysis of hyperspectral data," *Remote Sensing Environ.*, vol. 66, pp. 41-51, 1998.

[7] R.L. Huguenin and J.L. Jones, "Intelligent Information extraction from reflectance spectra: absorption band positions," *J. Geophys. Res.*, vol. 91, no. B9, pp. 9585-98, 1986.

[8] D.B. Nash and J.E. Conel, "Spectral reflectance systematics for mixtures of powdered hypersthene, labradorite, and ilmenite," *J. Geophys. Res.*, vol. 79, pp. 1615-1621, 1974.

[9] R.B. Singer and T.B. McCord, "Mars: Large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance," in *Proc. 10th Lunar Planetary Science Conf.*, 1979, pp. 1835-1848.

[10] D.A. Roberts, M.O. Smith, and J.B. Adams, "Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data," *Remote Sensing of Environment*, vol. 44, pp. 255-269, 1993.

[11] B. Hapke, "Bidirectional reflectance spectroscopy 1. Theory," *J. Geophys. Res.*, vol. 86, pp. 3039-3054, 1981.

[12] M.M. Verstraete, B. Pinty, and R.E. Dickinson, "A physical model of the bidirectional reflectance of vegetation canopies. 1. Theory," *J. Geophys. Res.*, vol. 95, no. D8, pp. 11755-11765, July 1990.

[13] J. Iaquinta, B. Pinty, and J.L. Privette, "Inversion of a physically based bidirectional reflectance model of vegetation," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 687-698, May 1997.

[14] G.P. Asner, A.R. Townsend, and B.H. Braswell, "Satellite observation of El Nino effects on Amazon forest phenology and productivity," *Geophys. Res. Lett.*, vol. 27, no. 7, pp. 981-984, Apr. 1, 2000.

[15] P.E. Johnson, M.O. Smith, S. Taylor-George, and J.B. Adams, "A semiempirical method for analysis of the reflectance spectra of binary mineral mixtures," *J. Geophys. Res.*, vol. 88, pp. 3557-3561, 1983.

[16] P.E. Johnson, M.O. Smith, and J.B. Adams, "Simple algorithms for remote determination of mineral abundances and article sizes from reflectance spectra," *J. Geophys. Res.*, vol. 97, pp. 2649-2658, 1992.

[17] J.F. Mustard and C.M. Pieters, "Quantitative abundance estimates from bidirectional reflectance measurements," in *Proc. 17th Lunar Planetary Science Conf.*, 1987, pp. E617-E626.

[18] J.F. Mustard and C.M. Pieters, "Photometric phase functions of common geologic minerals and applications to quantitative analysis of mineral mixture reflectance spectra," *J. Geophys. Res.*, vol. 94, pp. 13619-13634, 1989.

[19] J.F. Mustard and C. M. Pieters, "Abundance and distribution of ultramafic microbreccia in moses rock dike: quantitative application of mapping spectrometer data," *J. Geophys. Res.*, vol. 92, pp. 10376-10390, 1987.

[20] J.F. Mustard, L. Li, and G. He, "Nonlinear spectral mixture modeling of lunar multispectral data: Implications for lateral transport," *J. Geophys. Res.-Planets*, vol. 103, no. 19, pp. 419-425, 1998.

[21] L. Li and J. F. Mustard, "Compositional gradients across mare-highland contacts: The importance and geological implications of lateral mixing," *J. Geophys. Res.-Planets*, vol. 105, pp. 20431-20450, 2000.

[22] N. Keshava, J. Kerekes, D. Manolakis, and G. Shaw, "An algorithm taxonomy for hyperspectral unmixing," in *Proc. SPIE, Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI*, vol. 4049, 2000, pp. 42-63.

[23] B.-C. Gao, K.B. Heidebrecht, and A.F.H. Goetz, *Atmosphere Removal Program (ATREM) Version 2.0 Users Guide*. Boulder, CO: Center for the Study of Earth from Space/CIRES, University of Colorado, 1996.

[24] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sensing*, vol. 26, pp. 65-74, 1988.

[25] J.B. Lee, S. Woodyatt, and M. Berman, "Enhancement of high spectral resolution remote-sensing data by a noise-adjusted principal components transform," *IEEE Trans. Geosci. Remote Sensing*, vol. 28, pp. 295-304, 1990.

[26] J. Bowles, J. Antoniades, M. Baumback, J. Grossman, D. Haas, P. Palmadesso, and J. Stracka, "Real time analysis of hyperspectral data sets using NRL's ORASIS algorithm," *Proc. SPIE*, pp. 38-45, May 1997.

[27] J.B. Adams, M. O. Smith, and A. R. Gillispie, "Imaging spectroscopy: interpretations based on spectral mixture analysis." in *Remote Geochemical Analysis: Elemental and Mineralogical Composition*, C. M. Pieters and P. A. Englert, Eds. Cambridge, U.K.: Univ. of Cambridge, 1993, pp. 145-166.

[28] G. Foody and D Cox, "Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership model and fuzzy membership functions," *Int. J. Remote Sensing*, vol. 15, pp. 619-631, 1994.

[29] J.C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Comput. Geosoci.*, vol. 10, no. 2-3, pp. 191-203, 1984.

[30] A.D. Stocker and P. Schaum, "Application of stochastic mixing models to hyperspectral detection problems," in *Proc. SPIE, Algorithms for Multispectral and Hyperspectral Imagery III,* vol. 3071, Orlando, FL, Apr. 22-23, 1997, pp. 47-60.

[31] S. Tompkins, J.F. Mustard, C.M. Pieters, and D.W. Forsyth, "Optimization of endmembers for spectral mixture analysis," *Remote Sensing Environ.*, pp. 472-489, 1997.

[32] J.W. Boardman, "Geometric mixture analysis of imaging spectrometry data," in *Proc. Int. Conf. Geoscience and Remote Sensing*, vol. 4, 1994, pp. 2369-2371.

[33] M.D. Craig, "Minimum-volume transforms for remotely sensed data," *IEEE Trans. Geosci. Remote Sensing*, vol. 32, pp. 99-109, 1994.

[34] G. Strang, *Linear Algebra and Its Applications*. New York: Harcourt Brace Jovanovich, 1988.

[35] S.M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ: Prentice Hall, 1993.

[36] C.L. Lawson and R.J. Hanson, *Solving Least Squares Problems*. Englewood Cliffs, NJ: Prentice-Hall, 1974.

[37] D.G. Manolakis, V.K. Ingle, and S.M. Kogon, *Statistical and Adaptive Signal Processing*. New York: McGraw-Hill, 2000.

[38] M.S. Ramsey and P. R. Christensen, "Mineral abundance determination: Quantitative deconvolution of thermal emission spectra," *J. Geophys. Res.*, vol. 103, no. B1, pp. 577-96, Jan. 1998.

[39] D.A. Roberts, M. Gardner, R. Church, S. Ustin, G. Scheer,and R.O. Green, "Mapping chapparal in the Santa Monica mountains using multiple endmember spectral mixture models," *Remote Sensing Environ.*, vol. 65, pp. 267-279, 1998.

[40] S.G. Herzog and J.F. Mustard, *Lunar and Planetary Science XXVII*, pp. 535-536, 1996.