Chapter 5

Spectral Analysis for Earth Science: Investigations Using Remote Sensing Data

John F. Mustard
Brown University
Providence, Rhode Island

Jessica M. Sunshine
Science Applications International Corporation
Chantilly, Virginia

5.1 INTRODUCTION

Remote sensing in the earth sciences has grown remarkably over the last decade. This growth has been spurred by rapid advances in technology that have opened up new avenues of analysis and inquiry. Of these advances, the merging of spectroscopy and imaging has been the most important. Spectroscopy has been used as a quantitative tool in the laboratory for many years and there exists a wealth of understanding and analysis strategies for such data. Although early imaging spectrometer instruments suffered through the usual development problems, these systems are now approaching the spectral resolution and quality of laboratory measurements, blurring the distinction between the two but also bringing some of the more advanced laboratory spectral analysis techniques to bear on complex earth science problems of the field. The most advanced sensors and instruments are currently mounted on aircraft, but there are exciting plans to integrate the best of these into orbiting platforms which will facilitate greater accessibility and wider geographic coverage.

Broadly speaking, spectral analysis refers to the extraction of quantitative or qual-
itative information from reflectance spectra based on the albedo- and wavelength-dependent reflectance properties of materials. The goal of this chapter is to highlight some of the major developments in the use of spectral analysis for earth science applications that have evolved over the last decade, since the preceding edition of the Manual of Remote Sensing. Many of these developments have been pushed forward through the efforts of NASA and the Jet Propulsion Laboratory in the development of imaging spectroscopy (e.g., Goetz et al., 1985), although other agencies and companies [notably Geophysical Environmental Research (GER) and the Canadian Centre for Remote Sensing] have also contributed substantially in this area. The preceding Manual of Remote Sensing contained many well-documented and excellent examples in the use of broadband data analysis techniques (broadband refers to sensing systems that have coarse, low spectral resolution, like the Landsat thematic mapper). We do not devote much effort to these areas, except to identify the major approaches and the theoretical and spectroscopic bases for their success. The majority of this chapter is devoted to the use of spectroscopic techniques for material identification and quantification.

The foundation for quantitative analysis of remote sensing data through spectral analysis is reflectance spectroscopy. In general, reflectance is defined as the ratio of the intensity of the electromagnetic radiation scattered from a surface to the intensity of the radiation incident upon it. When measured as a function of wavelength, reflectance spectra exhibit specific albedo, continuum, and absorption features which are a function of the material properties of the surface measured. The absorption features are related to the chemical composition and mineralogy of the surface, while the continuum and overall albedo are a function of nonselective absorption and scattering as well as broad wavelength selective absorptions. These spectral properties, which cover broad wavelength regions, define the continuum of a spectrum and are partially controlled by both the physical properties of the surface (particle size, roughness, texture, etc.) and the chemical composition.

The physical principles governing reflectance spectroscopy have been elucidated through many theoretical, laboratory, and observational studies over the last three decades. In Chapter 1, Clark provides an excellent review of these principles with examples that illustrate the reflectance properties of materials with different mineralogy, composition, and texture. The goal of spectral analysis is to exploit this understanding and modify it for the analysis of remotely sensed data in order to achieve specific research and/or applied science objectives. The specific objectives may range from mapping of broad geologic units in support of reconnaissance geologic investigation to the detailed analysis of spectral signatures to map compositional variations related to a process. Despite what may appear to be a radically different approaches and methodologies to data analysis, all such investigations exploit the wavelength-dependent properties of materials as outlined by Clark in Chapter 1.

Presently, there is no single, universally accepted methodology for spectral analysis of remote sensing data. This is due in part to the enormous growth in the use of remote sensing for earth science investigations over the last two decades and a parallel growth in the number of techniques for data analysis. Frequently, the specific methodology used in a given investigation is optimized for the local objectives. The most effective techniques, and those that are more readily transported from one field site or data set to another, are those with solid foundations in the basic principles of reflectance spectroscopy. Within the diversity of techniques, we can generalize three
basic categories of spectral analysis: definition and mapping of broad-scale units, identification of the presence of specific mineralogic assemblages or lithologic units, and quantification of the amount of material present. The techniques for spectral analysis progress from simple techniques, suitable for discriminating different units but not necessarily providing compositional determinations, to precise deterministic techniques that require high spectral resolution and fidelity as well as a solid foundation in reflectance spectroscopy.

5.1.1 Definition and Mapping of Broad-Scale Units

A basic requirement of many earth science investigations is the existence of maps of surface compositions. The creation of a map involves simplification of small-scale complexity through the identification of similarities and trends and organization of this information into coherent and physically meaningful units. Units are the fundamental means of recording the properties of the surface and conveying this information to others. From the earliest investigations using aerial photography, the definition of broad-scale units was recognized as one the strengths of remote sensing. With the evolution of technology from black-and-white photographs to digital multispectral images acquired from orbit, these capabilities have been widely recognized and utilized. Units are defined on the basis of shared textural and spectral properties as exhibited in the multispectral data. Although some underlying physical basis can be identified, for the most part the specific compositional properties of the units cannot be determined uniquely. Field investigations for verification and ground truthing can be used to assign field-based names to units defined on the basis of remote sensing analyses, which can then be extended throughout the investigation site or across regions. For the most part, the definition and mapping of broad-scale units can be accomplished with broadband sensors through the use of simple spectral analysis techniques (e.g., color composites, band ratios, principal components analysis).

5.1.2 Identification of Specific Minerals, Assemblages, or Lithologic Units

The identification of specific mineral and/or lithologic compositions requires spectral analysis techniques that are based on the principles of reflectance spectroscopy. When the number of spectral channels is small, such as in multispectral systems, determination of broad mineral and surface compositional classes is possible. For example, with broadband sensors it is possible to identify regions likely to be enriched in ferric minerals and to distinguish these from areas enriched in hydrous minerals or vegetated regions. However, due to the lack of spectral resolution, it is not possible to define the unit composition. When the number of spectral channels of an instrument is sufficiently high (often loosely referred to as high spectral resolution) and spectral coverage spans a wavelength range that includes many individual absorption bands such that the spectral properties are well defined (e.g. absorption band shape, strength, etc.), then a determination of mineral type and composition, mineral assemblage, and lithology is possible. Typically, the methods for detailed characterization require ground-truth information in the form of a spectral library. The spectral
library must be representative of the material in the field and may include laboratory and/or in situ spectra of materials in the field.

### 5.1.3 Quantification of the Amount of Material

For many investigations, identification of materials present in a scene and the quantification of their abundance in a spatial context is critical and important information. Mixtures of materials within the field of view of remote sensing instruments is a natural phenomenon, due both to the physical size of the image pixel and because earth processes tend to create mixed surfaces. Weathering of rocks, for example, creates rinds of altered minerals surrounding fresher rock in the interior. Over time, the rind breaks off and is transported away, exposing more fresh rock. Thus rock outcrops may contain mixtures of altered and unaltered minerals at the surface. The transported material may end up in soils, which contain minerals transported from elsewhere combined with those formed in situ and with organic material. As another example, vegetation is tightly coupled to rainfall in semiarid regions. Where there are strong environmental gradients such as in rain shadows or mountain fronts, continuous changes in vegetation abundance occur with no distinct boundaries separating classes or zones. Such situations have been notoriously difficult to characterize and analyze with classification algorithms but are highly amenable to constituent abundance algorithms.

### 5.2 INFORMATION CONTENT OF REMOTE SENSING DATA

Effective use of spectral analysis for earth science investigations must begin with basic considerations of the overall objectives of a project and the potential information content of the remotely sensed data. These two categories are intimately coupled and careful early planning is an important step in optimizing the value of the results within available resources. It may be tempting to think that because higher spatial and spectral resolution will increase your ability to resolve objects spatially and discriminate their composition, the best approach is to obtain the highest resolution possible over an entire field site. However, data with high spatial and spectral resolution increases the need for human and computational resources for data processing and analysis. For the highest resolutions available, these costs are currently prohibitive if the analysis includes wide area coverage.

Definitions of spatial, spectral, and radiometric resolutions vary considerably because the usage and definitions have evolved from different technical and applications-oriented communities. It is certainly not the intent of this chapter to consider and explain fully all the possible definitions, but it is important to establish the basic meanings for the terms that have been adopted for this discussion. Spatial resolution in earth science applications typically refers to the ability to identify objects and/or determine patterns and spatial arrangements in imaging data. This notion of spatial resolution is a complex concept to define, as it varies with the instantaneous field of view of the detector (IFOV), the Raleigh criterion, spectral contrast in the scene, and
so on (e.g., Simonett et al., 1983). However, for simplicity, we adopt the definition based on the geometric properties of the imaging system. The smallest resolution cell in an image is defined by the IFOV and is typically quoted as the pixel size in meters or the angular resolution of the detectors in milliradians. Spectral resolution is determined by the bandwidth of the channels but is also coupled to the spectral coverage of the system. Narrow bandwidths permit a better characterization of the spectral signatures of materials, as long as the spectral coverage includes diagnostic absorptions features of the objects of interest. For example, 10-nm bandwidth channels between 0.4 and 1.1 µm would be effective for characterization and discrimination of ferric minerals, but the same wavelength range would be useless for investigations concerned with clay and/or carbonate minerals which have diagnostic absorptions in the wavelength region 2 to 2.5 µm. Radiometric resolution refers to the maximum number of discrete levels available for encoding reflected or emitted radiation. For a typical 8-bit system this corresponds to 256 levels (e.g., Landsat), while a 12-bit systems would have a maximum of 4096 levels. Although it is desirable to acquire data with the greatest number of levels, and thus a finer radiometric resolution, there are practical limits related to instrument fidelity, data acquisition and transmission, and storage.

The practical trade-offs between spectral and spatial resolution can be illustrated with a few examples. Consider the objectives for reconnaissance mapping of a large region (e.g., several hundred kilometers on a side or \( \geq 10^5 \) km\(^2\)) for mineral exploration that include the discrimination of rock outcrops from areas of thick soil cover or obscuring vegetation, and a first-order understanding of the distribution of major lithologic units. These objectives may be characterized as defining broad-scale units and mapping their distributions. If remotely sensed data with 1-m spatial resolution and 100 spectral channels were acquired for this area, on the order of \( 10^{12} \) bytes of data would need to be processed. Given the basic objectives, this investment is probably not warranted. On the other hand, a more modest data acquisition plan involving a system with 50-m spatial resolution and 10 spectral channels would result in a more manageable data volume of \( \geq 10^8 \) bytes, and provided that the 10 spectral channels are well placed for discrimination of the basic units, sufficient spectral dimensionality to achieve the basic objectives. As another scenario, consider the need to map alteration assemblages of clay and ferric minerals in a well-exposed outcrop of a small region (e.g., 10 km\(^2\)). This would require high spectral resolution data (\( \geq 10 \) nm/spectral channels) to distinguish between the different ferric and clay mineralogies. The small size of the region would permit a high spatial resolution as well.

This trade-off between spectral and spatial resolution affects the design and operation of many current and planned remote sensing systems (see Chapter 11) and is well illustrated by the moderate-resolution imaging radiometer (MODIS) (Salomonson and Toll, 1991; King et al., 1992). The MODIS instrument will be used to measure and monitor biological and physical properties, including surface temperature, ocean color, the condition and dynamics of vegetation and land surface covers, and cloud and aerosol properties. This instrument can provide data in up to 36 spectral bands (20 between 0.4 and 3.0 µm and 16 between 3 and 15 µm), and it is designed to obtain global coverage every 1 to 2 days. Although it will have the capability to obtain a maximum spatial resolution of 250 m per pixel at nadir, the large swath width (2300 km), coupled with numerous bands and a finite data rate, precludes the use of this high-resolution mode for the collection of global data sets.
Instead, the instrument can be programmed to maximize spectral and spatial resolution of the data returned for the specific application required. When a greater spectral resolution is required to identify and map compositionally dependent parameters, data would be acquired with a lower spatial resolution. For applications that require higher spatial resolution, the key physical properties will need to be resolved with fewer bands.

5.3 SIMPLE METHODS OF SPECTRAL ANALYSIS

Spectral analysis need not be complicated in order to extract important information from a remotely sensed data set or to achieve a specific set of objectives. All that is really required is that the materials of interest can be discriminated objectively. A number of enhancement and digital processing techniques have been developed with this goal in mind. When applied with a concrete understanding of the physical causes of absorption (see Chapter 1), simple methods are semiquantitative and can be very effective tools. Most of these methods are used in the analysis of broadband sensors much as the Landsat thematic mapper (TM), which lack the spectral resolution and coverage required for the more sophisticated and deterministic approaches (described later in this chapter). Nevertheless, the approaches can also be quite effective for systems with higher spectral resolution.

5.3.1 Color Composites

Color composite images are the most basic form of spectral analysis. The selection of specific spectral bands for display in the red, green, and blue image planes of a computer display can be used for rapid first-order analysis of the spectral properties of a scene even in the absence of calibration (whenever possible, 24-bit displays should be used, as significant information may be lost in compressing the full range of color in a three-band color composite to an 8-bit display). When the image data are stretched to maximize the dynamic range displayed, the relative color variations of the surface materials become evident.

As an example, consider the images shown in Figure 5.1 see color insert). These are two color composites of a Landsat TM scene over a portion of the Grand Canyon, Arizona. In the top image, the data are displayed in "natural" color, which approximates how the terrain would appear to the human eye [TM band 1 (0.45 to 0.52 μm) in blue, 2 (0.52 to 0.60 μm) in green, and 3 (0.63 to 0.69 μm) in red]. The earthy tones are certainly familiar colors of typical rocks in this region. In contrast, the bottom image shows the same region but uses two of the infrared channels with Landsat TM band 3 (0.63 to 0.69 μm) in blue, 4 (0.76 to 0.90 μm) in green, and 7 (2.08 to 2.35 μm) in red. The very different view actually can be used for rapid assessment of vegetated versus nonvegetated regions as well as certain rock properties. In this color composite, vegetated areas appear green since the reflectance properties of vegetation are a maximum in TM band 4, a minimum in TM band 3, and typically low in TM band 7 (although this depends on the water content of the vegetation). Soils and rocks with abundance ferric oxides exhibit low reflectances in band 3, with moderate to high reflectances in bands 4 and 7, and regions rich in
these minerals exhibit the distinctive yellow colors. The rocks in the lower parts of the canyon have higher amounts of amphiboles, which tend to have lower reflectances in the visible to near-infrared, but exhibit increases in reflectance in band 7 and thus appear reddish. There are also some lithologies in these scenes with high reflectances in the visible but declining toward longer wavelengths and appear in colors of blue and cyan. By selecting various combinations of bands, different spectral properties can be emphasized and/or suppressed, which when combined with a basic knowledge of the spectral properties of materials expected in a scene, can provide an excellent first-order assessment of the diversity of units in a region.

An objective method for determining the optimum bands for display has been developed by Crippen (1989). A correlation matrix for the suite of bands in a given scene is calculated. The correlation coefficients for each pairwise channel combination is then used to calculate a ranking index:

$$\text{index} = \sqrt{1 + 2abc - a^2 - b^2 - c^2}$$  \hspace{1cm} (5.1)

where $a$, $b$, and $c$ are the pairwise correlation coefficients. The maximum index value of 1.0 indicates the three bands that display the maximum spectral variability and an index of 0.0 indicates the minimum. This approach can also be used with band ratios or spectral indices (described below). Although the index provides an objective method to determine the best three bands for maximum spectral diversity in a color composite display, these may not be the best for enhancing the discrimination of materials of interest. For example, the maximum spectral diversity may be controlled by vegetation abundance, while the analysis objective is lithologic discrimination.

### 5.3.2 Band Ratios

Ratio images, created by dividing one spectral image channel by another pixel by pixel, are widely used in earth science investigations because of the proven ability of ratio images to enhance the discrimination of surface units. When applied with a well-formulated rationale and spectroscopic basis, they can be very effective spectral analysis tools. The principal strengths of ratio images are that the effects of topographic shading are removed, albedo differences are suppressed, and the relative color properties of materials are enhanced. Before calculating a ratio image, the data must be corrected for any system contribution (e.g., dark current) or environmental signal (e.g., atmospheric path radiance) that is additive to the digital numbers (DNs) measured. These additive signals contribute significantly to the calculation of the ratios and will cause erroneous results unless removed.

This is well illustrated in Figure 5.2 (see color insert), which is the same region as that shown in Figure 1 and uses the same spectral bands. The band ratio images and color composite of the ratios shown in Figure 5.2a were created using the raw digital data, with no corrections for scattering. The consequences of this negligent calibration can readily be identified in ratio images. Except for regions dominated by materials with extraordinary spectral contrast (e.g., vegetation), the resultant images are correlated to the albedo of the input data, and topographic shading is still apparent in the ratio image. Note that this effect is enhanced for the ratios using band 1, which has the greatest amount of atmospheric scattering. Compare these with the band
ratio images shown in Figure 5.2b, where the data were corrected for atmospheric scattering. Not only are the effects of topographic shading and albedo eliminated from the product, there are dramatic differences in spatial distribution of color units. It is now easier to trace like-color compositional units around the canyon walls, but also highlights the distinct color properties of the topmost rock unit from the vegetated areas of the plateau. Multiplicative corrections would also be required if the absolute value of the ratios was important information, but are not necessary if the objectives are simply to discriminate surface units.

Band ratios have been most used extensively in the analysis of broadband sensors such as Landsat multispectral scanner (MSS), TM, and SPOT, spurred on by the successes of many of the early studies (Rowan et al., 1974; Goetz et al., 1975; Ashley and Abrams, 1980; Rowan and Kahle, 1982; Powdysocki et al., 1983). Ratios can also be used for the analysis of data from narrowband imaging spectrometer systems, although the continuous spectral coverage permits more sophisticated methods (see below). The most common band ratios exploit two major classes of absorption: visible to near-infrared absorptions due to ferric and ferrous-bearing minerals and infrared absorptions associated with hydrous, hydroxyl, and carbonate minerals. The spectra shown in Figure 5.3a indicate that ferric and ferrous-bearing minerals exhibit diagnostic absorptions in the visible to near-infrared wavelength regions (see Chapter 1). There are distinct differences in spectral properties of the minerals hematite and goethite relative to the nonferric minerals kaolinite and montmorillonite are a consequence of the electronic transition and charge transfer absorptions that occur in the ferric compounds (Sherman and Waite, 1985; Chapter 1, this volume). The same data filtered to the spectral bandpasses for the TM sensor are shown in Figure 5.3b. While the diagnostic features that distinguish kaolinite from montmorillonite are not present, the major differences between the ferric minerals and the hydroxyl-bearing minerals are evident.

The common multispectral ratios for the minerals shown in Figure 5.3 are shown in Table 5.1. TM band 3 is positioned near the first maximum in the spectra of the ferric oxides, and thus the ratio of TM band 3 to band 1 will highlight the strong increase in reflectance from the ultraviolet through the visible. This ratio for materials lacking ferric compounds will be close to unity (e.g., montmorillonite and kaolinite, Table 5.1), while for ferric-bearing compounds the ratio will be much greater than 1 (e.g., hematite and goethite, Table 5.1). In addition, many ferric minerals exhibit an increase in reflectance between 0.9 and 1.7 \( \mu m \). Thus the ratio of TM band 5 to band 4 will be much greater than 1 for surfaces enriched in these compounds and near or less than unity for other materials (Table 5.1).

With the wavelength coverage of the TM system, it is also possible to highlight the presence of such hydrous minerals as kaolinite, montmorillonite, alunite, and others. These minerals typically show a decrease in reflectance from 1.7 to 2.2 \( \mu m \) due to overtones and combination overtones of water and hydroxyl fundamental absorptions (Chapter 1) (Figure 5.3a). Therefore, the ratio of TM band 5 to band 7 increases with increasing amounts of these minerals and/or the combined strength of these absorption features (Table 5.1). There are many possible band ratio combinations and the specific combinations used should be constructed with reference to the spectral properties of the materials present in the field site. For example, Sultan et al. (1987) used several ratios of TM bands that were based on analysis of the field and laboratory reflectance spectra of lithologies from the Arabian-Nubian Shield in
5.3 Simple Methods of Spectral Analysis

Egypt. They showed that the ratio of TM band 5 to band 1 was highly correlated with the volume percent of opaques, $5/7$ with hydroxyl-bearing phases, and the product of two ratios, $3/4$ and $5/4$, with ferrous absorptions in amphiboles. Multiple-band-ratio images can be used to construct color composites that allow the simultaneous analysis of several features. Additional examples of applications using band ratios are presented by Sabine in Chapter 8.

Ratios provide a simple and rapid tool for analyzing spectral properties of surface materials and discriminating different units. When ratios are applied with a good knowledge of the spectral properties of the key materials in the scene, they can pro-
vide useful first-order information about the surface composition. Ratios do not have a sufficient level of sophistication to make deterministic statements of surface composition, largely because materials may exhibit similar ratio values but have different spectral and compositional causes. However, some spectral features are so strong and characteristic of specific materials (e.g., the red edge in vegetation) that ratios can be used to identify areas with these properties rapidly and therefore can be effective for subsetting data.

### 5.3.3 Statistical Transforms

Ratio techniques are not effective for discriminating among materials that exhibit more subtle spectral differences. An inherent property of multichannel remote sensing data is that the spectral properties of materials are highly correlated channel to channel and across much of the wavelength region. This is a consequence of the general properties of reflectances of solid surface. Reflectance, which is constrained to lie between 0 and 1.0, changes intensity slowly from channel to channel. Although the different channels are often thought of as independent variables, they are not; there are no negative values, and spectral discontinuities, where reflectances exhibit large changes from one channel to the next, are very rare. Vegetation, which typically transitions from less than 10% reflectance at 0.6 μm to greater than 60% at 0.8 μm, is one exception, although even this change in reflectance occurs over a 200-nm wavelength interval. The discrimination of different lithologies and soils is commonly based on very subtle, but consistent and real, spectral differences. Even band ratios, where the fundamental albedo information has been removed, show high degrees of correlation.

A number of approaches have been developed to maximize the spectral variance of remote sensing data to assist in the discrimination of different surface materials. The most widely used and well developed of these techniques is principal components analysis (PCA) (e.g., Taylor, 1974; Williams, 1983; Loughlin, 1991). In PCA the original data are projected onto a new set of orthogonal coordinates that are defined by the statistical properties of the input data. PCA requires no a priori information and, unlike band ratios or color composites, is based on the entire spectral domain of the data set. Because the new output images are linear, additive combinations of the original input images and depend on the spectral variability within the scene analyzed, the individual images are not readily interpretable in terms of specific absorptions or minerals as are band ratios. However, through the use of ground truth and evaluation of the weightings given to each band in the principal components, it

### TABLE 5.1 TM Band Ratios for Spectra Shown in Figure 5.3

<table>
<thead>
<tr>
<th>Mineral</th>
<th>3/1</th>
<th>5/4</th>
<th>5/7</th>
<th>3/4×5/4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Montmorillonite</td>
<td>1.1262</td>
<td>0.95793</td>
<td>0.61064</td>
<td>0.60173</td>
</tr>
<tr>
<td>Kaolinite</td>
<td>1.1262</td>
<td>1.0033</td>
<td>0.67157</td>
<td>0.65574</td>
</tr>
<tr>
<td>Hematite</td>
<td>7.2598</td>
<td>1.8767</td>
<td>0.07343</td>
<td>0.06211</td>
</tr>
<tr>
<td>Goethite</td>
<td>3.8719</td>
<td>1.9934</td>
<td>0.14474</td>
<td>0.13756</td>
</tr>
</tbody>
</table>
is possible to build more quantitative understanding into this approach (e.g., Smith et al., 1985; Jaumann, 1991; Huete, 1986).

PCA is generally based on the correlation matrix between spectral channels, which is constructed with data from the entire scene, a representative sample of the scene, or a subset of the scene. The correlation matrix is decomposed into its characteristic eigenvalues and eigenvectors using standard matrix decomposition routines, the details of which will not be covered here. Conceptually, this decomposition defines a new set of axes that are linear combinations of the original variables (spectral channels in this case) that maximize the amount of variance that each axis describes and are orthogonal to one another. The components are ordered where the first axis (or principal component) describes the most variance and the last component the least. The amount of the spectral variability contained in each component is given by the eigenvalue, and the relative proportion or contribution of each band to that component is given by the eigenvector. When reflectance or raw radiance data are input, the first principal component is typically a weighted average of all the bands used and is similar in appearance to an albedo image of the scene. When vegetation comprises a significant component in a scene, it is commonly highlighted in the second principal component, which is a consequence of the strong spectral contrast between the visible and near-infrared reflectance spectra of vegetation. The next components typically highlight spectral variability due to rocks and soils.

In broadband systems such as TM, the latter principal components are dominated by random and systemic noise and are usually of little interest for interpretation and analysis. With hyperspectral and imaging spectrometer data, however, the latter components may contain important information. It is usually worthwhile inspecting these results because statistical outliers which comprise only a small proportion of the scene's spectral variance may be contained in these components. Green et al. (1988) have developed a very useful modification of PCA called the minimum noise fraction (MNF) MNF transform. An estimate of the noise covariance matrix is first derived, and then the reflectance data are rotated and scaled to make the noise isotropic with unit variance in all bands. These transformed data are then analyzed using principal components. Because the noise of the data set is prewhitened and distributed equally among the bands, this approach results in a better discrimination of the eigenvalues related to surface-spectral elements from those related to instrument and scene noise. This better orders the large number of components produced with imaging spectrometer data to put spectrally and spatially relevant components first.

PCA may be implemented with a more quantitative spectral basis through a number of different strategies. One method proposed by Crósta and Moore (1989) and Loughlin (1991) uses subsets of the total number of possible spectral channels. The particular bands are chosen to maximize the variance due to a set of absorption features. For TM images, a PCA is performed separately on band combinations 1, 3, 4, 5 and 1, 4, 5, 7. The first set of bands emphasizes spectral variability in the visible regions, which is coupled to ferric and ferrous iron absorptions. The second set emphasizes spectral variability at longer wavelengths and is coupled to hydroxyl absorptions. Each analysis maximizes the variance in the input data for a given suite of absorptions and thus allows a separation of materials based on these different spectral properties. When the full TM band set is used, the spectral variability from these different wavelength regions is combined. PCA has also been applied to data sets with more spectral channels, and in these applications it is possible to interpret
the eigenvectors in terms of specific spectral properties and features. For example, Jaumann (1991) performed PCA on a suite of rock and soil laboratory spectra and showed that when plotted as a function of wavelength, the first few eigenvectors, exhibited features that were analogous to mineral absorption bands seen in the spectra of the rocks and soils. He then correlated the eigenvectors to the major element oxide chemistry of the samples using multiple linear regression. These eigenvectors and regression coefficients were then used to map surface composition for a remotely acquired data set. In this example, ground-truth data were used to construct a PCA framework that was then applied to the remote data. Therefore, the remote observations were transformed into a PCA data space defined by laboratory spectra.

Smith et al. (1985) and Pieters et al. (1985) used PCA to model mixing relationships between spectroscopic end members. This approach is predicated on the assumption that the spectral variability contained in the data set is due fundamentally to mixing between a small number of discrete compositional endmembers with unique spectral properties. Other approaches to mixing are discussed in detail below. If the mixing systematics are linear, or can be linearized, the relative distance between end members as measured in the transformed PCA data space is proportional to the abundance of the end members. In a similar approach, Huete (1986) decomposed the spectral variability of laboratory data from a large number of soils into the principal components and showed that the first four eigenvectors, combined linearly, accounted for greater than 98% of the spectral variability of the soil suite. By this approach it is possible to account for much of the variability in diverse soil spectra with a small number (four in this case) of basis eigenvectors. Each of the basis vectors was shown to be correlated to specific soil types, the spectra of which are a function of the composition.

PCA and other statistical transforms are thus effective tools for summarizing the common components of spectral variability of a scene into a small number of variables. The separation of materials in a remote sensing scene by PCA is driven by the spectral variability of the materials and is thus linked to spectral properties. The principal limitation of this approach is that the statistical transforms are scene specific, and therefore the results and methodologies for interpretation are not easily transported to different regions or to scenes acquired during different seasons. However, through the use of laboratory data to define the PCA space (e.g., Smith et al., 1985; Jaumann, 1991; Huete, 1986) or subsetting the data into specific spectral regions (e.g., Crósta and Moore, 1989; Loughlin, 1991), it is possible to bring a stronger measure of spectral understanding to the analysis of data with statistical transforms.

5.4 FEATURE MAPPING AND THE IDENTIFICATION OF MINERALS

The identification and mapping of specific minerals is one of the great strengths of imaging spectroscopy. This capability is important not only for mapping the presence of minerals over wide areas, but also because it can be used to discriminate among minerals that may not be readily identified through inspection in the field. An important distinction of this type of approach from those presented in the preceding discussion is that feature mapping and mineral identification is physically based. It
5.4 Feature Mapping and the Identification of Minerals

exploits the fact that many minerals exhibit absorption bands that are diagnostic of mineral type and composition (the physical basis for absorption and the relationship to mineral structure and chemistry is presented in Chapter 1). By using as much of the available spectral information as possible, absorption features can be quantitatively characterized and then compared and analyzed in the context of spectral libraries (e.g., JPL, USGS, RELAB), thus linking the remote measurements to a form of ground truth. Because the basic physical interactions in laboratory and remote measurements of spectral properties are the same, this is a valid approach, although there are clearly issues of scale and texture that must be considered. It should be obvious that an explicit requirement for this type of technique is that the remote measurements must be calibrated to reflectance.

There are three basic approaches to mapping and analyzing mineral absorption features. The first simply characterizes the position, strength, and shape of absorption features. The results can then be displayed as absorption feature maps or used as input to analytical systems to compare the results to laboratory spectra characterized in the same way. The second method seeks to compare the complete shape of absorption features from a remote data set to those in a spectral library or database. This allows the direct identification of minerals and combinations of minerals or assemblages. The third approach, which attempts to deconvolve overlapping and superimposed absorptions quantitatively, has applications in the identification of mineral combinations and abundances and for mineral compositions across solid solutions.

5.4.1 Position, Strength, and Shape

If presented with the reflectance spectrum of a mineral, an experienced analyst is able to identify quickly the particular mineral on the basis of the absorption bands present in the spectrum. The key features that are used in this process are the position, shape, and relative strength of all the absorptions present. Most minerals exhibit a number of distinct absorptions that vary in position, shape, and strength, which in combination are a unique indicator of the particular mineral species (e.g., Burns, 1993; Hunt and Salisbury, 1970; Hunt et al., 1973; Chapter 1, this volume). Implicit in this approach is that there are wavelength regions that do not exhibit absorptions and that the absorption features are in essence departures from a smoothly varying and continuous function. This continuous function is commonly referred to as the continuum.

The continuum of spectrum is poorly understood and difficult to define. It may be qualitatively defined as the collective properties of spectral regions exhibiting smoothly varying spectral properties that, taken as a whole, define the upward limit of the general reflectance curve for a material. Like all reflectance data, the shape and properties of a continuum includes both absorption and scattering processes. However, unlike absorption features, the absorption processes are generally nonselective or extremely broad and ill defined (Clark and Roush, 1984). Scattering processes also contribute significantly to the continuum, perhaps best illustrated by the effects of changing particle size on reflectance spectra (e.g., Nash and Conel, 1974; Pieters, 1983; Crown and Pieters, 1987; Chapter 1, this volume). The solid surface of a freshly exposed rock generally exhibits a low reflectance and relatively weak
absorption bands. As this rock is broken up into fine particles, the overall reflectance brightens, the absorption bands become deeper, and the continuum becomes accentuated. Further decreases in particle size result in a weakening of the absorptions due to the increase in scattering relative to volume absorption (Hapke, 1993). Except for the transition from a coherent rock to particulate, the continuum generally scales multiplicatively with textural changes and illumination and thus is a relatively fixed or constant feature of a material.

Full and quantitative analysis of absorption features for mineralogic identification currently requires detailed examination and modeling of individual spectra using optical constants and complex scattering codes (Chapter 1). Due to the large size of hyperspectral data sets and the lack of high-quality optical constants for the vast majority of earth materials, this approach is not feasible for terrestrial data sets, and more automated methods that are robust and stable are desirable. A number of methods have been developed over the last decade to automate these procedures (e.g., Green and Craig, 1985; Kruse et al., 1986, 1988; Yamaguchi and Lyon, 1986). The method of Kruse et al. (1988) is well tested and has been integrated into expert systems and analytical packages (e.g., Kruse et al., 1993). This work forms the basis of the following discussion. However, the other techniques are based on the same principles and differ primarily in the methods of implementation.

The basic concepts of this approach are illustrated in Figure 5.4. The initial step in the analysis is the definition of the continuum. This can be achieved automatically by defining the high points in the spectrum through the use of slope and magnitude criteria. A high point is defined simply by the wavelength where the reflectance is higher than the reflectance values at short and longer wavelengths. The additional criterion that the slope is positive on the short-wavelength side and negative on the long-wavelength side of this point over a given wavelength range helps to minimize the selection of spurious points due to noise. It is prudent to use a smoothing function to suppress the effects of noise and calibration error (Kruse et al., 1988). Straight-line segments are drawn between the defined high points and then the actual reflectance and that defined by the straight-line segments are compared for each channel in the data set. The reflectances for the model continuum are the larger of the two compared values. This is considered a model continuum since it is an approximation based on the data and not derived from first principles using optical constants and scattering models. A number of investigators have developed successful approaches to define the continuum, which is actually an important step in many spectral analysis techniques (e.g., Green and Craig, 1985; Clark et al., 1987).

The model continuum approximates the convex hull for the spectrum. The continuum is then removed from the reflectance data by dividing the reflectance by the continuum at each channel (Figure 5.4). The objective of the continuum removal is to isolate the properties of the absorption features from the overall reflectance properties of other components in the signal. The overall reflectance can be well modeled as exponential functions of the absorption coefficient and mean optical path length of each component multiplied together (Clark and Roush, 1984; Hapke, 1993). Therefore, to preserve and properly scale the critical strength, shape, and position parameters, the reflectance must be divided by the continuum. This provides a first-order correction for the effects of illumination that is also a multiplicative factor.

The continuum-removed data are then used to define the wavelength position of the minima, strengths, and asymmetries of discrete absorption bands (Figure 5.4).
Figure 5.4 Schematic diagram illustrating the basic concepts behind mapping absorption band position, strength, and shape using a laboratory spectrum of calcite as an example. (a) definition of the continuum; (b) spectral curve after the continuum is removed; the band position is defined as the lowest point in the relative reflectance spectrum, the depth by equation (5.1), and the full width at half maximum (FWHM) as illustrated; (c) definition of asymmetry for negative, zero, and positive asymmetry. Reprinted from Remote Sensing of Environment, Vol. 44, F. A. Kruse, A. B. Lekhoff, and J. B. Dietz, Expert system-based mapping in northern Death Valley, California/Nevada using the airborne visible/infrared imaging spectrometer (AVIRIS), pp. 309-335, Copyright 1993, with permission from Elsevier Science.
Band minima are identified in an analogous way to the methods used in determining the high points for the continuum, except it is the low points that are defined and cataloged. If there are more than two band minima identified between a given pair of high points, the absorption band is defined as a multiple band and assigned an order of two or greater, depending on the number of minima encountered. The relative strength of each band is calculated using the formula presented by Clark and Roush (1984):

\[ S_1 = \frac{R_e - R_b}{R_e} \quad \text{or} \quad S_2 = 1.0 - \frac{R_e - R_b}{R_e} \]  

(5.2)

where \( S \) is the band strength, \( R_e \) the reflectance of the continuum at the wavelength of \( R_b \), and \( R_b \) the reflectance at the band minimum. For \( S_1 \), band strength is normalized to be between 0.0 and 1.0, where weak bands have a strength near 1.0 and strong bands have values much less than 1.0. \( S_2 \), also normalized to be between 0 and 1.0, is essentially the inverse of \( S_1 \), and weak bands have strengths close to 0.

Band shape can be parameterized for the purposes of this systematic analysis by two simple parameters: full width at half maximum and asymmetry (Figure 5.4). The width of the feature is the absolute difference in microns between the right and left sides of the absorption, where the reflectances are half the strength of the band. Asymmetry is defined as the base 10 logarithm of the sum of reflectances over the number of channels to the right of the band minimum divided by the sum of reflectances over the number of channels to the left of the band minimum. The use of the \( \log_{10} \) preserves the linearity in calculated asymmetry for bands with equal asymmetries to the right and left of the band minimum. The result is that a perfectly symmetric band will have a value of 0, asymmetric to the right will be positive and to the left, negative.

This parameterization allows a rapid reduction in the dimensionality of an imaging spectrometer data set to a suite of key absorption features. The specific parameters for the example shown in Figure 5.4 are presented in Table 5.2. This reduces the spectral data to four discrete absorptions, each parameterized by four variables. Although this is a useful summary of the number and character of absorptions, this information in and of itself does not necessarily provide a link to the surface mineralogy. To make practical use of the results from such an analysis, they must be coupled to ground truth, either through field or laboratory spectra. Kruse et al. (1993) have integrated the band parameterization approach with spectral library information into an expert system for analysis of hyperspectral data. The basic method involves reducing the library data to a table of band parameters and then

### Table 5.2 Absorption Band Parameters for the Mineral Calcite from Figure 5.4

<table>
<thead>
<tr>
<th>Band</th>
<th>Wavelength</th>
<th>Depth</th>
<th>FWHM</th>
<th>Asymmetry</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.340</td>
<td>0.3001</td>
<td>0.0980</td>
<td>0.3047</td>
</tr>
<tr>
<td>2</td>
<td>1.997</td>
<td>0.0788</td>
<td>0.0490</td>
<td>0.5154</td>
</tr>
<tr>
<td>3</td>
<td>1.870</td>
<td>0.0603</td>
<td>0.0490</td>
<td>0.6033</td>
</tr>
<tr>
<td>4</td>
<td>2.164</td>
<td>0.0452</td>
<td>0.0392</td>
<td>0.1947</td>
</tr>
</tbody>
</table>
using these results to build a hierarchy of facts and rules to catalog the remote data. These decisions cascade from first-order discrimination of vegetated from nonvegetated surfaces, rocks from soils, hydrated minerals from ferric minerals, and so on. Kruse et al. (1993) analyzed the probabilities for 27 different minerals with this approach in a data set for the northern Grapevine Mountains, Nevada, and mapped with reasonable accuracy the distribution of five specific minerals. Although this approach requires high signal/noise ratio (>50:1), it can be an effective starting point in an analysis.

5.4.2 Feature Mapping Using Complete Band Shape

The method described above is rapid and effective for processing the large amounts of data typical of hyperspectral data sets. When coupled to complementary analyses of spectral libraries, it can yield reasonably accurate results for the identification of minerals and their spatial distributions. However, the approach above is sensitive to noise, which limits the applicability. In addition, the method for band parameterization condenses into the four variables the amount of spectral information from a given band. As such, much information concerning the specifics of the band shape are lost from the analysis, which can lead to ambiguity in identifying minerals. Although less of a concern for narrow strong absorptions with simple structure, this becomes a greater problem for absorptions with greater width (e.g., ferric and ferrous absorptions) and complexity (multiple hydroxyl bands with differences in concavity, etc.). Furthermore, the spectral contrast of absorptions in remotely acquired data sets is generally lower than for laboratory or field spectra. This is primarily a consequence of mixing of components at the subpixel scale, and therefore simple comparisons of band strength from remotely acquired spectra to spectral libraries will be complicated.

Clark and others (Clark et al., 1990, Clark and Swayze, 1995) have developed a technique for absorption band mapping that uses the entire suite of spectral information within a given spectral range to define the band. This approach is intimately coupled to data from a spectral library, since the algorithm computes the degree of similarity between the remote data and the library data. A continuum is first removed from both library and remote data using predefined channels on either side of the absorption feature (Figure 5.5a). Averaging of several proximal channels is generally performed to increase the signal/noise ratio for these points. Thus a priori knowledge is a requirement since the analysis must begin with a determination of specific minerals or components that are to be mapped or analyzed in the remote data set.

The continuum is removed from each spectrum around the absorption band by dividing the reflectance by the straight-line continuum, resulting in a new suite of spectra, defined as

\[ L_c(\lambda) = \frac{L(\lambda)}{C(\lambda)} \quad \text{and} \quad O_c(\lambda) = \frac{O(\lambda)}{C_0(\lambda)} \] (5.3)

where \( L(\lambda) \) is the library spectrum as a function of wavelength \( (\lambda) \), \( C(\lambda) \) is the continuum for the library spectrum, \( L_c \) the continuum removed library spectrum, \( O \) and \( O_c \) the
observed and continuum removed observed spectra, respectively, and $C_0$ is the continuum for the observed spectrum.

The next step in this approach is to adjust the contrast of the continuum removed library spectrum [$L_c(\lambda)$] to best fit the continuum removed observed spectrum [$O_c(\lambda)$] using a simple linear gain and offset adjustment. This is not strictly valid from a theoretical standpoint, as it does not correctly account for band saturation at low albedos or large grain sizes, nor does it properly account for nonlinear affects of intimate mixing on band strength and shape. However, most of the variance due to these effects can be simulated to a first order with this approximation, and the ap-
The approach is computationally efficient, thus allowing rapid analysis of large volumes of data. A new continuum-removed reference library spectrum, \( L'_c \), is calculated by using an additive constant, \( k \), in the equation

\[
L'_c = \frac{L_c + k}{1.0 + k}
\]  

(5.4)

which can be rewritten into the standard form of the linear equation:

\[
L'_c = a + bL_c
\]  

(5.5)

The constants \( a \) and \( b \) are defined in terms of \( k \) by the equations

\[
a = \frac{k}{1.0 + k} \quad \text{and} \quad b = \frac{1.0}{1.0 + k}
\]  

(5.6)

This can be rearranged to a standard least-squares equation where the values of \( a \) and \( b \) are determined that best match the spectrum of the continuum-removed reference spectrum, \( L_c(\lambda) \), to the continuum-removed observed spectrum, \( O_c(\lambda) \):

\[
L'_c = a + bL_c
\]  

(5.7)

Since, by definition, the number of spectral channels will be greater than or equal to three (two to define the continuum, one for the band center), this is an overdetermined problem and a solution will always exist. This results in a superposition of the library spectrum by the linear gain and offset into the same reference frame as the observed spectrum (Figure 5.5b). The parameters that can then be derived are the band strength as defined in equation (5.2), and the root-mean-square (RMS) difference between the remote and library spectra. The RMS calculation provides an overall goodness-of-fit measure, while the band depth is proportional to the abundance of the mineral. In an automated analysis of image data, the goodness of fit can be set to a predetermined threshold level, depending on the noise in the observed data set, where an RMS value above this threshold would indicate the presence of the given mineral. These RMS and band depth measures can be combined by multiplying them together to produce a map of the distribution and relative spectral abundance of minerals. Ancillary information such as albedo and continuum properties can also be included in the algorithm. For example, when searching for the mineral magnetite, the algorithm also checks that the albedo is generally low, or if the target material is vegetation, the continuum should have a steep rise from the visible to near-infrared. This additional information permits increased fidelity in discriminating among materials.

An excellent example of this approach is presented by Clark et al. (1995), where they have used this method to map, in detail, the presence of a wide variety of minerals at Cuprite, Nevada (Figure 5.6). Cuprite has been the object of numerous remote sensing studies, beginning with analysis of data from broadband sensors by Ashley and Abrams (1980) and followed by a number of hyperspectral studies (e.g., Goetz and Srivastava, 1985; Hook and Rast, 1990; Kruse et al., 1990; Swayze et al.,
Figure 5.6 Alteration map of Cuprite, Nevada. This shows the basic alteration assemblages for this well-studied geologic site and forms the base map for the results shown in Figure 5.7. A wide variety of alteration minerals have been identified at this site, including quartz, alunite, kaolinite, buddingtonite, and ferric oxides. (After Ashley and Abrams, 1980; Hook and Rast, 1990.)

1992). The well-documented surface mineralogy, excellent lithologic exposures, sparse vegetation cover, and limited soil development have contributed to the selection of Cuprite as a test site for evaluation of sensor and algorithm performance. The geology of the site is characterized by acid sulfate hydrothermal systems in two discrete centers. The eastern center consists of hydrothermal alteration of Tertiary volcanic rocks, while in the western center, Cambrian metasedimentary rocks have been altered. The alteration zones typically follow a concentric pattern from weakly altered argillized rocks on the outside through opalized to silicified rocks in the interior. The extraordinary array of well-exposed lithologies and minerals in this deposit is a consequence of the rapidly changing fluid compositions and temperatures during its formation in what is interpreted to have been a hot spring deposit (Hook and Rast, 1990; Buchanan, 1981).

The image in Figure 5.7a (see the color insert) provides an overview of Cuprite in simulated TM true-color bandpasses. The AVIRIS data have been convolved to approximate the spectral band passes of the TM sensor, and the scene is approximately 17 km long and 10.5 km wide. The mineral map shown in Figure 5.7b (color
5.4 Feature Mapping and the Identification of Minerals

insert) is derived from analysis of AVIRIS data, emphasizing minerals with relatively narrow vibrational absorptions in the spectral region 2 to 2.5 \( \mu \)m (i.e., \( \text{OH}^- \), \( \text{CO}_3^- \), and \( \text{SO}_4^- \) bearing minerals; see Chapter 1). The reference mineral spectra are taken from the USGS spectral library, and the map indicates areas in the AVIRIS scene that contain absorptions that match those selected from the spectral library. Regions in black indicate an absence of these minerals. The mineral map shown in Figure 5.7c (color insert) is derived from analysis of the same data but with a focus on electronic absorption features associated with \( \text{Fe}^{2+} \) and \( \text{Fe}^{3+} \) bearing minerals, such as hematite, goethite, and amphiboles. These absorption features are broader than the vibrational absorptions and are typically found in the spectral region 0.4 to 1.2 \( \mu \)m. These two figures show 41 different mineral categories out of more than 250 specific materials that were mapped in this scene (although not all materials were present). By combining the two mineral maps (b and c) it is possible to construct detailed mineral assemblage maps. This is a tremendous achievement in discriminating among minerals and surface materials with remote sensing and begins to demonstrate the level information available from high-quality imaging spectroscopy data sets.

Consider a little more carefully the results shown in Figure 5.7. In Figure 5.7b, Clark et al. (1995) have been able to resolve subtle shifts in absorption positions related to solid solution series and element substitutions and mapped minerals with these characteristics (e.g. high, medium, and low aluminum muscovite; K-alunites of different temperatures). This level of mineralogic discrimination requires excellent radiometric resolution, precision, and accuracy as well as a thorough spectral library. The discrimination of dickite from kaolinite is a good example of this, as they have very similar absorptions (Chapter 1). In Figure 5.7c the primary focus is on minerals with electronic absorption features associated. Again, solid solution substitutions in \( \text{Fe}^{2+} \) and \( \text{Fe}^{3+} \) bearing minerals affect the absorption band properties, as does the degree of crystallinity of the minerals (e.g., large, medium, fine, and nanophase hematite). This field site has been analyzed extensively over the years, and the mineral maps presented here have been confirmed through field and laboratory studies (Swayze et al., 1992, Swayze and Clark, 1995). The results shown in Figure 5.7 are a dramatic example of the high level of precision in mineral mapping that is possible with high-quality imaging spectrometer data (in this case, 1995 AVIRIS data; see Chapter 11) and well-characterized library spectra.

5.4.3 Absorption Band Modeling

The approaches to absorption feature mapping described above rely on correlations between the parameterization of features observed in remotely acquired data to those in spectral libraries or ground-truth data sets. These methods are therefore excellent for applications for which spectral libraries representative of the materials in the scene exist. They also include a certain degree of flexibility to handle noisy data and complexities introduced by mixtures or particle size effects. However, when presented with mineralogic compositions that are different from those found in spectral libraries and/or with several minerals in a mixed surface, simple matching of absorption features can prove to be inadequate. Mineral identification is particularly difficult when overlapping absorptions combine, obscuring the minima of one or more of the diagnostic absorption bands. Models that account more completely for
the processes of absorption by focusing on individual absorption bands rather then combined absorption features are more easily extended beyond the scope of preexisting libraries.

One absorption band-based approach, developed by Huguenin and Jones (1986), examines the various higher-order derivatives of spectra to identify the location of individual absorptions. This method assumes that each absorption is symmetric around its band center, but unlike other approaches discussed below, does not require that absorptions have a specific shape. Band centers are identified where the second derivative of the spectrum is negative, the fourth derivative is positive, and the fifth derivative is zero. Like any derivative analysis, this method is highly sensitive to noise. Therefore, the Huguenin and Jones approach is critically dependent on its incorporation of an intelligent smoothing algorithm, the details of which remain proprietary to the authors. Nonetheless, the approach is capable of resolving overlapping band centers separated by as little as 0.3 to 1.0 of the full width at half maximum (assuming Gaussian shaped absorptions). Results derived with this approach are very complementary to that of Kruse et al. (1988) and Clark et al. (1990).

Many workers (e.g., Burns, 1970; Smith and Strens, 1976; Farr et al., 1980; Clark, 1981; Singer, 1981) have approximated absorption bands with Gaussian distributions of the form

\[ g(x) = \exp \left[ -\frac{(x - \mu)^2}{2\sigma^2} \right] \]  

(5.8)

where \( x \) is wavenumber (energy), \( s \) the strength or amplitude of the distribution, \( \mu \) the center (mean) wavenumber, and \( \sigma \) the width. However, the simple Gaussian approximation has been shown to be an inadequate physical model for absorption bands (Sunshine et al., 1990). This is particularly well illustrated for the absorption bands in the mineral orthopyroxene. The 0.9-\( \mu \)m absorption is due to electronic transitions from a single site in the orthopyroxene structure (Burns, 1970) and thus should be modeled by a single Gaussian, if the absorption is in fact Gaussian in form. However, two Gaussian distributions are required to fit the feature adequately (e.g., Singer, 1981; Roush and Singer, 1986).

The underlying assumption of the Gaussian model is that energy (wavenumber) is the random variable. However, the center of a band, the location of the maximum probability of absorption, is a function of the average crystallographic and molecular properties of the material. Similarly, the width of absorptions is primarily a consequence of perturbations to these average properties. Fundamentally, the energy of absorption is determined by the average bond length of the atoms defining the absorption site. Thus, to the first order the random variable should be variations in bond length (Sunshine et al., 1990).

Since the average bond length is randomly or Gaussian distributed, the key determining factor for the shape of absorption is how bond-length variations are mapped into wavelength or energy. According to the crystal field theory for electronic absorptions (Marfunin, 1979; Burns, 1993), absorption energy (e.g., wavelength), and bond length are related by a power law. Sunshine et al. (1990) empirically determined the optimum mapping of the bond length to energy by fitting an absorption due to a single absorption, the 0.9-\( \mu \)m orthopyroxene absorption. This analysis revealed
that energy and bond length are inversely proportional, suggesting that Coulombic potential energy dominates the absorption site. This inverse mapping of bond lengths to energy leads to a modified Gaussian distribution, \( m(x) \), of the form

\[
m(x) = \text{sexp} \left[ -\frac{(x^{-1} - \mu^{-1})^2}{2\sigma^2} \right]
\]  

(5.9)

The difference between Gaussian and modified Gaussian distribution is reflected in the symmetry of the band models.

Having formulated a physically based mathematical description for isolated absorption bands, Sunshine et al. (1990) were able to model spectra accurately as a series of modified Gaussian distributions superimposed on a continuum. In the modified Gaussian model (MGM), the continuum is modeled as a straight line in energy (cm\(^{-1}\)) that sits above the general spectral curve. It should be noted that continua used in most other methods (e.g., Kruse et al., 1988; Clark et al., 1990) use reflectance maxima to define the continuum. This difference is critically important for the case of overlapping absorptions, where the true reflectance maxima are obscured. Under the MGM each absorption is characterized by three model parameters: a band center (\( \mu \)), a bandwidth (\( \sigma \)), and a band strength (\( s \)), while the continuum is characterized by the slope and offset of a straight line in energy. The model calculations are performed in natural log reflectance and energy, and all parameters, including the continuum, are fit simultaneously using a stochastic least-squares method (Tarantola and Valette, 1982).

Examples of MGM solutions to a low-calcium pyroxene, a high-calcium pyroxene, and mixtures of the two minerals are shown in Figure 5.8. Each panel in this figure includes (top to bottom) the residual error between log of the modeled spectrum and log of the actual spectrum (offset 10%), the individual modified Gaussian distributions representing absorption bands, the continuum, and the modeled spectrum superimposed on the actual spectrum. [The nonrandom residual error near 0.9 \( \mu \)m in (a) is symmetric with respect to the modeled absorption. Unlike an asymmetric residual, which generally indicates the presence of an additional absorption band, this symmetric error is a characteristic of minor band saturation, as discussed by Sunshine and Pieters (1993).] Each of the pure minerals has a unique MGM fit to the spectrum, where the strength, center, and width of the absorptions are characteristic of that mineral and controlled by its chemistry and crystal structure. In the mixed spectra it is not immediately apparent that there are two mineral phases, as there is only one band minimum in both the 1-and 2-\( \mu \)m regions. Yet, using the MGM, it is possible to deconvolve these overlapping absorptions successfully. Furthermore, modified Gaussian distributions from the solutions to the mixed spectra match in position and width the modified Gaussians calculated for the pure minerals. The strengths of the absorptions vary in proportion to the abundance of each mineral in the mixture.

Sunshine and Pieters (1993) analyzed absorption band systematics in pyroxene mixtures using the MGM and derived a particle-size-independent relationship between the relative band strengths of the 1-and 2-\( \mu \)m bands and the abundance of these minerals in the mixtures. Sunshine et al. (1993) subsequently applied this relationship to spectra of natural samples that contained variable amounts of pyrox-
Figure 5.8 Modified Gaussian model (MGM) fits of reflectance spectra: (a) MGM fit to orthopyroxene size separate; (b) MGM fit to clinopyroxene size separate; (c) MGM fit to 75% orthopyroxene–25% clinopyroxene mass fraction mixture; (d) MGM fit to 25% orthopyroxene–75% clinopyroxene mixture. In (c) and (d) open arrows correspond to absorptions from the low-Ca orthopyroxene phase. Solid arrows correspond to absorptions from the high-Ca clinopyroxene.
enes and other minerals. They were able to deconvolve successfully the mafic mineral abundances in these more complex natural lithologies. The MGM has also been used to model the variations in absorption band properties in solid solution series (Sunshine, 1994; Mustard, 1992). In these applications it was shown that the modified Gaussian distributions vary systematically in position and strength with mineral composition and that relationships could be derived that related the absorption band properties to composition. The study of Sunshine (1994) examined the changes in electronic transition absorptions across the olivine solid solution series. Mustard (1992) focused on the actinolite solid solution series, where absorptions due to electronic transitions, as well as charge transfers and vibrations, were present. Although the MGM model was not developed specifically to model hydroxyl or charge transfer absorptions, the general principles behind the MGM model apply. Therefore, for minerals that exhibit systematic changes in vibrational absorptions as a function of composition and/or environment of formation, the MGM may be an effective tool for characterizing these absorptions and relating them to composition and environment of formation.

It should be noted that in contrast to other methods described above, the MGM neither relies on a library of spectra that may or may not reproduce all permutations and combinations of absorption bands, nor requires knowledge of end-member spectra, as is needed for both linear and nonlinear mixing described below. Instead, the MGM derives compositional information directly from a measured spectrum. This is particularly useful in applications where the exact mineralogy, major and minor element chemistry, and physical state (particle size, alteration state, etc.) of materials is unknown or may not exist in library spectra.

5.5 Full Spectral Mapping

As discussed above, mapping minerals based on the absorption features in their spectra is quite powerful. Although such methods are strongly linked to composition, they are by definition focused only on specific wavelength regions and therefore neglect contributions from the rest of the spectrum. Feature mapping is particularly problematic for many materials (e.g., rocks), whose spectra do not contain well-defined absorptions. Instead of diagnostic features, these spectra are characterized by their continuum shapes and/or very broad absorptions. To address some of these concerns, a number of alternative and complementary approaches have been developed that compare spectra simultaneously over their full wavelength range.

There are two general classes of approaches to mapping based on the entire spectral signature. The first group of methods are spectral similarity searches. Under a spectral similarity search, a scene is examined to determine which pixels are most "similar" to a specific spectrum of interest or reference spectra. As discussed below, there are numerous ways of objectively defining the similarity between two spectra, all of which can be used to produce spectral similarity maps. A second category of mapping methods that uses the full spectral response are spectral detectors or matched filters. Here a spectrum of interest is detected in a scene by highlighting pixels with similar spectral properties simultaneously while repressing all other background spectral signatures. Spectral detection maps can be produced by many ap-
approaches, each of which is based on a different characterization of the spectrum of interest and the background.

### 5.5.1 Spectral Similarity Maps

One of the fundamental uses of imaging spectroscopy is to determine the spatial extent of specific materials based on their spectral response. For example, one application of interest might be to remotely map the location of a particular outcrop. This can be accomplished by first identifying a representative spectrum of the outcrop from either a known location within the scene or from field data and/or a spectral library. (The use of spectra from field instruments or a library does, of course, require that the scene be calibrated to absolute reflectance and that the reference spectrum be resampled to the same wavelengths as the remote data.) The scene can then be searched pixel by pixel for spectra that are, based on some predefined metric, most like the representative outcrop spectrum. A similarity image is thus generated, with high values being most similar to the outcrop spectrum. Typically, a threshold value must then be set to define a tolerance level for what is considered similar enough (within the noise) to be mapped as the outcrop. There are a variety of algorithms that have been used to define the similarity between two spectra. As discussed below, the diversity in approaches stem from the fact that a spectrum can be viewed from many different perspectives in addition to its physical and/or compositional meaning.

One approach to defining similarity is to compare the difference between two spectra at each wavelength mathematically. Although a variety of metrics could be imagined, it is common to use the RMS distance:

\[
\text{RMS} = \sqrt{\frac{\sum_{i=1}^{N} (R_i - P_i)^2}{N}}
\]

(5.10)

where \(N\) is the number of bands in the spectrum, \(R\), the value of the spectrum of interest or reference spectrum at the \(i\)th band, and \(P\), the value of the spectrum of a pixel of the scene at the \(i\)th band. The RMS difference can be calculated for each pixel, with low RMS values indicating areas that are most similar to the spectrum of interest. To produce a map, some maximum threshold RMS value would be chosen. This RMS distance (or RMS error) is often used a measurement of goodness of fit (e.g., Clark and Swayze, 1995, as discussed above) and is a good choice for defining overall spectral similarity from a statistical perspective.

An alternative viewpoint is to consider spectra geometrically. The spectral response as a function of wavelength, while compositionally significant, can also be interpreted geometrically as a vector in \(N\)-space (where \(N\) is the number of bands). Visualizing spectra as vectors leads naturally to at least two metrics for defining similarity. One approach is to use the vector dot product as a similarity measurement. The dot product between two vectors is

\[
R_i \cdot P_i = \sum_{i=1}^{N} R_i P_i
\]

(5.11)
The dot product is proportional to the projection of one vector onto the other. Thus pixels with large vector projections onto the spectrum of interest could be considered most similar. After the appropriate thresholding, the dot product could then be used to produce a similarity map.

A complementary vector-based approach is to measure the angular distance between two vectors. Under the so-called spectral angle mapper (e.g., Kruse et al., 1993), spectra whose vectors are separated by small angles are considered most similar. By measuring angular distance and ignoring vector length, this metric downplays differences in albedo and as such provides a good complement to the dot product approach.

One other common similarity method for spectral searching is binary encoding. Binary encoding is based on Hammond distances (e.g., Mazer et al., 1988) and comes from the field of digital communications. Under this algorithm, each spectrum or signal is first divided by its average. At each wavelength, if the spectrum is greater than or equal to its average, it is assigned a value of 1. If it is below the spectral average, a value of 0 is assigned. Thus, based on variations in amplitude, a generalized description of the spectrum is encoded into a 1 bit vector (each point or wavelength is either 0 or 1). To preserve additional information on the shape of the spectrum, local spectral slope is also encoded. The local slope is defined at each wavelength, \( i \), as simply

\[
R_i = \frac{R_{i+1} - R_{i-1}}{2}
\]

The second bit is set to 1 if the slope is positive and 0 if the slope is negative. After encoding spectra into this 2-bit scheme, a measurement of similarity known as the Hamming distance can be used. The Hamming distance is defined as

\[
\sum_{i=1}^{N} B(R)_i \text{ XOR } B(P)_i
\]

where \( N \) is the number of bands in the spectrum, \( B(R)_i \), the binary-encoded reference spectrum (or spectrum of interest) at the \( i \)th band, \( B(P)_i \), the binary-encoded spectrum of a pixel of the scene at the \( i \)th band, and XOR the the exclusive OR logical operator. At each wavelength \( i \), the exclusive OR will return a value of 1 if the 2-bit spectra differ and a value of 0 if they are the same. After summing over all bands, the pixels with small Hamming distances (i.e., closest to 0) will be most similar to the reference spectrum. It should be noted that by defining amplitude relative to the mean (to encode the first bit) the binary encoding process is insensitive to albedo differences between spectra. Some variations in the details of the binary encoding method exist. In particular, some implementations use 1-bit encoding based only on the amplitude or only on the local slopes information. The binary encoding of the spectra allows for a very rapid calculation of the XOR logical operation, and as such, this similarity metric is extremely computationally efficient. An excellent application of the binary encoding technique for mapping minerals in the Grapevine Mountains can be found in Kruse et al. (1993).

Each of these similarity measurements in and of themselves is reasonable, straightforward to implement, and can be used to produce a similarity image. However,
there are several issues of general concern. First, there is no clear choice as to which method to use. Furthermore, the various methods, each derived from a different perspective, can produce different results. As none of the approaches are strongly compositionally based, there is, in general, no reason to give more weight to any particular method. In some cases, however, a technique may be better suited for a particular scenario. For example, neither binary encoding nor SAM retain albedo information. In situations where albedo is meaningful, the RMS difference may be a better metric. While the choice of methods remains uncertain, it should be noted that obtaining similar results from different methods does provide a greater degree of confidence in the solution. In all cases it is clear that careful evaluation of results from several techniques is warranted. A second concern is that each of the approaches requires that an acceptance threshold be defined in order to produce a map showing the spatial extent of a given material (e.g., an outcrop). While some combination of experience with the method and the particular scene can aid in establishing a reasonable threshold, determining the threshold value remains somewhat arbitrary. As such, care must be taken to examine the sensitivity of the resulting map to changes in the threshold value.

5.5.2 Spectral Detection Maps

The basis of spectral similarity mapping is spectral comparisons and similarity indexes. In contrast, spectral detection mapping centers around the concept of maximizing the signal of the spectrum of interest while suppressing all background spectra, or in the language of signal processing, maximizing the signal/clutter ratio. The key difference between similarity searches and spectral detection algorithms is that the latter accounts directly for the presence of background materials, while the former ignores them. Spectral detection, often referred to as matched filtering, can be thought of conceptually as a compromise between processes that maximize the contribution of the signal of interest (e.g., spectral similarity searches) and those that minimize the remaining background. Actual determination of the optimal matched filter depends on assumptions about the state of knowledge of both the spectrum of interest and its background. In some cases approaches require significant knowledge of the spectral properties of both [e.g., foreground–background analysis (FBA); Smith et al., 1994]. Alternatively, some techniques rely on statistical estimates from the scene or generalized spectral characteristics.

One of the more fully developed spectral detection algorithms was recently presented by Harsanyi and others (Harsanyi, 1993; Harsanyi and Chang, 1994). Their technique addresses the specific problem of identifying materials with known spectral responses that are present only at low abundances. Under Harsanyi's low-probability detection (LPD) algorithm, the background spectral signatures are assumed to be unknown and are statistically estimated from the scene. There are many situations where this theoretical framework is applicable, such as the case where a material exists only at a scale below the resolution of the sensor (i.e., has subpixel exposures). This could, for example, describe an outcrop that is either partially covered by vegetation or exposed only in small areas. Any measurement of subpixel exposures would, of course, include a spectral contribution not only from the material of in-
terest (e.g., the outcrop) but also from other scene components (e.g., vegetation) and is thus well suited to the LPD approach.

As with the similarity mapping techniques, the first step of the LPD is to define the spectral signature of interest from within the scene or from a spectral library. (In the case of a library spectrum, it must first be converted to at-sensor radiance or DN). Under the LPD, each pixel in the scene, \( P \), is then described as a linear combination of the reference spectrum (or spectrum of interest), \( R \), and the unknown background spectral properties, \( U \), plus system noise, \( n \):

\[
P = R \alpha + U \alpha + n \tag{5.14}
\]

where \( \alpha \) is the abundance of the reference spectrum and the vector \( \alpha \) is the abundance of unknown background spectra.

An orthogonal subspace projection, \( S \), is used to solve equation (5.14), where

\[
S = I - UU^* \tag{5.15}
\]

and \( I \) is the identity matrix and \( U^* \) is the pseudoinverse or \((U^TU)^{-1}U^T\). When equation (5.14) is multiplied by this orthogonal subspace projection operator, the contribution of the background, \( U \alpha \), is eliminated and the abundance of the reference spectrum can be determined.

Recall, however, that the background spectral properties, \( U \), are unknown. Using the scene itself, it is possible to estimate the background contribution statistically. This is accomplished by calculating the data correlation matrix:

\[
C = \frac{1}{p} \sum_{i=1}^{p} P_i P_i^T \tag{5.16}
\]

where \( p \) is the number of pixels in the scene and \( P_i \) is the spectrum of the \( i \)th pixel. The correlation matrix can then be decomposed into a matrix of eigenvectors, \( V \), and their associated eigenvalues using a principal component transform, as described in previous sections. As in PCA analysis, the first \( N \) eigenvectors can be chosen to be significant. As such, the first \( N \) eigenvectors, \( V_N \), can be used as a statistical approximation of the background. \( V_N \) can then be substituted for the unknown background contribution \( U \) and the orthogonal subspace projection, equation (5.15), can be rewritten as

\[
S = (I - V_NV_N^*) \tag{5.17}
\]

Having fully defined the LPD procedure, the definition of low probability is clearer. For the LPD approach to work, the spectrum of interest must be a minor scene component in the sense such that it is not included among the primary eigenvectors (the first \( N \)) used to estimated the background. As discussed previously, there is no one-to-one correspondence between principal components and spectral properties. Thus the statistical definitions inherently required by LPD may be difficult to comply with, as many materials that are minor components may statistically have
properties that are described at least partially by the primary eigenvectors. In addition, care must be taken in evaluating the output of the LPD, in that it may include pixels which have low-probability spectral properties (i.e., are not represented by the principal eigenvectors) but that are spectrally different from the reference spectrum of interest. It is clear, however, that for situations which conform to the theoretical framework of the LPD, the LPD detection scheme is quite powerful.

Several examples of the success of the LPD in discrimination of poorly exposed lithologies are documented by Farrand and Harsanyi (1995). One of the goals of this study was to map the spatial extent of palagonite tuffs found along the rim of Easy Chair Crater (ECC), part of the Lunar Crater Volcanic Field in Nevada, using imaging spectrometer data from a 1992 AVIRIS scene (Figure 5.9a). Spectra of representative lithologies of interest were collected from field samples and measured with the NASA-supported reflectance spectrometer facility at Brown University (RELAB) (Pieters, 1983; Mustard and Pieters, 1989) and are shown in Figure 5.10. The principal distinguishing spectral properties among the materials shown in Figure 5.10 are not specific absorption features but the overall continuum. Spatially, the ECC palagonite tuff is confined to a small region along the south and west rim of Easy Chair Crater (Figure 5.9). Spectrally, the ECC tuff is very similar to other background materials as shown in Figure 5.10. In fact, the overall continuum properties can be reasonably well modeled by a spectral mixture of the rhyolite and orange and red cinders. The spectral similarity of the ECC palagonite tuff to the background materials made discrimination by standard spectral mixture analysis problematic (spectral mixture analysis is described in detail below). Either the regions rich in the ECC palagonite tuff were adequately modeled by rhyolite and red cinder end members, or if an actual spectral end member for the tuff was used, it predicted high tuff abundances in regions where the tuff did not exist (Figure 5.9b).

However, the small spatial extent of the ECC palagonite tuff is well suited to the LPD method. Using the reference spectrum of the ECC palagonite tuff (after conversion to radiances units) as the known spectrum of interest, Farrand and Harsanyi (1995) used the LPD algorithm successfully to map the spatial extent of the ECC palagonite tuff (Figure 5.9c). The LPD isolates the ECC palagonite tuff dramatically while suppressing other scene components. The brightest values correspond directly to outcrops of the palagonite as defined by field work. The lower-intensity pixels correspond to smaller outcrops with low abundances of the tuff at the scale of the pixel.

### 5.5.3 Summary

The use of the full spectral range and resolution for identification and mapping of surface materials can be used to provide useful results. The methods discussed here range from spectral similarity techniques, which provide broad correlations, to the low-probability detection algorithm, which attempts to highlight areas with spectral properties similar to a target material with narrow spectral limits. It is important to note that all these approaches provide results, but they may not necessarily give the same results, because of the different algorithms used. When applying these methods of spectral analysis, it is useful to employ a combination of methods. Where there is a high degree of correlation between methods, a greater degree of confidence in the
5.6 Mixture Modeling

Material identification is obviously indicated. As with all the approaches to spectral analysis discussed here, results should be cross-referenced and linked to ground-truth information.

**5.6 Mixture Modeling**

Many of the approaches described above that exploit the spectral properties of materials to identify surface composition make the implicit assumption that there is a one-to-one correspondence between the observations and ground truth (i.e., labo-
Figure 5.10  Reflectance spectra of Easy Chair Crater palagonite tuff compared to various background materials: (a) laboratory reflectance of volcanic materials exposed at the field site of the Lunar Lake volcanic field; (b) laboratory spectrum of the target material, ECC palagonite tuff, compared with the AVIRIS spectral properties other common components in the scene.
ratory or field measurements of spectral properties of small samples are representative of surfaces measured remotely). The fact that at the scale of most observing systems, the pixel contains a mixture of materials is generally treated as a second-order effect. As a second-order effect, it is accommodated through adding complexity to the approaches, such as mineral combinations or simply considered a consequence of working with natural scenes. Certainly, the principal goal for mapping in the earth sciences is to simplify the complexity observed on the ground into a set of well-defined units that can then be organized and used to understand and analyze larger-scale associations. However, anyone with experience in field observations will recognize that continuous variations in the proportions of components are the norm when viewed at scales that approximate the typical spatial element of a sensor. These may simply be in the proportions of components that are of little interest to the analyst (e.g., vegetation abundance when mapping geology or variations in illumination), or they may comprise a critical signal that could be the main objective of an investigation (e.g., changing proportions of minerals due to a facies change in a sedimentary sequence).

The concept of the mixed pixel has certainly been recognized for many years (e.g., Horowitz et al., 1975; Jackson, 1983; Hue, et al., 1984). However, the idea that mixing at the subpixel scale is a natural consequence of earth processes and an inherent feature of remote sensing data sets that can be quantitatively exploited has only evolved over the last two decades. Among the researchers who have contributed to this effort (e.g., Mustard and Pieters, 1987a; Boardman and Goetz, 1991; Farrand, 1991; Li and Strahler, 1986), Adams and co-workers (Adams and Adams, 1984; Adams et al., 1986, 1993; Gillespie et al., 1990; Smith et al., 1990a,b; Sabol et al., 1992) have conducted some of the more structured analyses and multifaceted approaches to the problem.

The basic premise of mixture modeling is that within a given scene, the surface is dominated by a small number of common materials that have relatively constant spectral properties. If most of the spectral variability within the scene is a result of varying proportions of these common components (end members), it follows that the spectral variability captured by the remote sensing system can be modeled by mixtures of these components. It has been documented that the reflectance spectrum of a mixture is a systematic combination of the reflectance spectra of the components in the mixture (e.g., Nash and Conel, 1974). The systematics are basically linear if the components are arranged in spatially distinct patterns, analogous to the squares on a checkerboard (e.g., Singer and Mc Cord, 1979). In this case the scattering and absorption of electromagnetic radiation is dominated by a single component on the surface, and thus the spectrum of a mixed pixel is a linear combination of the end-member spectra weighted by the areal coverage of each end member in the pixel. If, however, the components of interest are in an intimate association, like sand grains of different composition in a beach deposit, the mixing systematics between these different components are nonlinear. The spectral properties of the different end members become convolved in this case, because the electromagnetic radiation interacts with more than one end member as it is multiply scattered in the surface.

The question of whether linear or nonlinear processes dominate the spectral signatures of mixed pixels is still an unresolved question. It probably depends on a number of factors and conditions of the scene. A discussion of the implications of nonlinear mixing is included near the end of this section. Nevertheless, the linear
approach has been demonstrated in numerous applications to be an insightful 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. With this as background, we discuss spectral mixture analysis in more detail in the following sections, with some examples.

5.6.1 Theoretical Framework

The principal objective of spectral mixture analysis is to define a coherent set of spectral end members that are representative of physical components on the surface and that model the spectral variability inherent in a given scene. The theoretical limit to the number of end members is defined by the dimensionality of the data and constraints of the mixture inversion. This results in a total number of possible end members equal to the number of spectral channels in the data set, plus 1. However, the number of end members that may be practically defined is far fewer, typically ranging from three to seven, depending on the number of channels and the spectral variability of the scene components, even for hyperspectral imaging spectrometer data sets. The reason for this is that the theoretical limit requires that the spectral channels be independent variables and that the spectral end members be linearly independent. 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 defined mathematically by linear combinations of other components. Such materials would therefore not be resolved. Strategies to overcome these limitations, such as multiple-end-member models, are discussed by Adams et al. (1993).

A practical way to approach the definition of end members is to apply a field-based framework in which classes of materials are considered (e.g., soils, vegetation, lithologies, shade) in the context of the basic technical constraints imposed by the data set to be examined (e.g., number of spectral channels, wavelength range, pixel size). It is also practical to begin with a few well-defined and spectrally distinct components, and build complexity into the model as is warranted by subsequent analyses. This is illustrated conceptually in Figure 5.11. Here 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 end-member shade 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 end member mixes with the other end members 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 end members (selected from the image data). An image end member (IE) is one that is contained within a scene and has a maximum abundance of the physical end member it is most
Figure 5.11 Conceptual diagrams illustrating the basic concept and terms of spectral mixture analysis: (a) visible–near infrared spectra of three end members—green vegetation, gabbro, and shade—at the spectral resolution of the thematic mapper sensor; (b) scatter plot of TM band 7 versus TM band 4 of data representative of a scene containing the variable proportions of the end members; image end members (IE) are represented by the open circles; (c) same scatter plot but with the reference end members plotted and the proper mixing space defined; (d) automated techniques for the identification of mixing end members can use the convex hull of the data cloud to define the vertices of the intersection where the reference end member should be (After Adams, 1993.)
closely associated with. This is illustrated in Figure 5.11b where the IE are represented by the open circles. They may be selected objectively (e.g., Adams et al., 1993; Boardman, 1993; Tompkins et al., 1997) 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 that corresponds to a pure component (the pixel is comprised of 100% of the end member), although some pixels will probably be relatively pure. Conceptually, the image end members are those that best bound the data cloud, as shown in Figure 5.11b. Once selected, these end members are used to solve the following equation for each pixel in the image:

$$DN_b = \sum_{i=1}^{N} F_i DN_{ib} + E_b \quad \text{and} \quad \sum_{i=1}^{N} F_i = 1.0$$

(5.18)

where $DN_b$ is the intensity of a given pixel in bandpass or wavelength $b$, $F_i$ the fractional abundance of end member $i$, $DN_{ib}$ the intensity of image end member $i$ at wavelength $b$, $N$ the number of end members, and $E_b$ the error of the fit for bandpass $b$. The second equation is a constraint that the fractions sum to 1.0. If $M$ is the number of channels in the data sets, this results in $M$ equations in $N$ unknowns and may be solved using standard linear least-squares inversion.

The fractions calculated for each of the data points (solid circles) shown in Figure 5.11b correspond to the relative distance from the image end member (IE) points, subject to the constraint that the sum of the fractions equals 1. Because the image end members are themselves mixtures of other materials, there will exist fractions in the results that will be greater than 1.0 or less than 0, although the sum of the fractions will still equal 1. Data points that fall within the lines joining the IEs would have positive fractions, while those that fall outside the lines would have negative fractions of the IE opposite this line. For example, those pixels that plot outside (to the right) of the line joining the image end members 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 application.

The validity of a mixture model solution using a particular suite of end members can be tested using three basic criteria: the root-mean-square (RMS) error of the fit of the end-member spectra to each of the pixels in the image, the spatial patterns and coherency of the fraction images, and the absolute values of the fractions. The RMS error is computed using the standard sum of the squares of the difference between the model and the data, normalized to the number of spectral channels. In general, an average RMS error is calculated for the entire scene as well as for each pixel, and this second measure can then be scaled and displayed for analysis (see the example below). If the average error is within the level of system noise, an adequate solution is indicated. The RMS error image can be analyzed for indications of where specific regions of the input data are not well modeled by the end members selected and these areas can be used to refine the selection of end members for subsequent analyses.

The fraction images can be displayed by scaling the fractions to fit the 8-bit dynamic range of typical computer displays [e.g., $100(F_i + 1)$ scales the fractions such that 0% equals a DN of 100 and 100% equals a DN of 200]. The spatial distribution and abundances of the end members represented in the fraction images should conform to a priori expectations or conventional spectral parameterizations for the scene. For example, the spatial patterns and abundance of vegetation should be similar to conventional vegetation indices (e.g., the NDVI), while the shade image should
be inversely correlated to the average albedo. Furthermore, the fraction images should have a certain level of coherency. A typical result when too many end members are used than is warranted by the data, or if two end members are used that are not spectrally distinct, is that the fraction images will exhibit little continuity in the fractions, with rapid changes between high and low values.

The absolute values of the fractions can be used to determine if the image end members selected properly bound the spectral variability of the data set. If there are significant regions of one or more fraction images that show either extreme positive or negative values, the end members are not well selected. Although the fractions will still sum to unity (that is a constraint in the original inversion), and the RMS error may be small, the specific image end members should nevertheless be reselected.

The results using image end members provide a first-order perspective on the mixing relationships, and if the end members are relatively pure, the fractional abundances will be similar to the actual abundances. However, it is often desirable to use library or reference end members to calculate the fractions. This provides a better link to spectral libraries and therefore ground truth, but is also essential if the fractional values are intended to be used in multitemporal studies or across scenes that were acquired under different illumination conditions or times of years. If the data are accurately calibrated to reflectance, the image end members can be related to the library reference spectra by

\[ DN_b = \sum_{i=1}^{N} F_i R_{ib} + E_b \quad \text{and} \quad \sum_{i=1}^{N} F_i = 1.0 \]  

(5.19)

where \( R_{ib} \) is the reflectance of library end member \( i \) in bandpass \( b \). This is illustrated in Figure 5.11c, where it can be seen that the spectral variability defined by the reference end members now fully bound the variability of the image data. In this example there would be no negative fractions or fractions greater than 1.0.

The primary result of the spectral mixture analysis is the fraction images which show the distribution and abundance of the end-member components in the scene. They can be used in a variety of ways to analyze the composition of the surface and provide insight into processes. Because the raw radiance data have been converted to data products that make intuitive sense from a field perspective, the fraction images can be used to identify lithologic units and map their distribution. The fraction images can be manipulated to suppress information extraneous to the task and enhance that which is more relevant. For example, the shade image is generally dominated by illumination effects that are not relevant to compositional mapping. The suite of fraction images can be normalized to a shade-free scene, thus allowing the abundance relationships among key components to be better illustrated. This would be accomplished by dividing each of the fraction images by \((1.0 - \text{shade})\). A similar procedure could be employed to remove the fractional contributions of vegetation to the scene, thus emphasizing rock and/or soil relationships.

The correspondence between the spectral abundance of materials determined from spectral mixture analysis and actual physical abundance on the ground depends on a number of factors. Smith et al. (1990b) have shown that green leaf abundance determined from leaf area index is highly correlated to the abundance of the green vegetation end member from spectral mixture analysis of arid regions, and Mustard and Pieters (1987b) have shown that mineral abundances from mixture modeling of hyperspectral data are well correlated with physical abundances in terrains where
the minerals exhibit unique and diagnostic absorptions. Sabol et al. (1992) have analyzed the general question of detectability for a variety of different materials, spectral imaging systems, and performance parameters. The detectability increases with the signal/noise ratio of the measurements, number and placement of spectral channels, and the spectral contrast of the materials. For high-spectral-resolution data and well-differentiated end members, the detectability may be very low, approaching a few percent, while for low-spectral-resolution systems with little spectral contrast among the end members, the detectability may be a few tens of percent. In a similar manner, Lawler and Adams (1994) performed an analysis of system and environmental noise for some typical measurement platforms and showed that formal uncertainties in the fractional abundances may be incorporated directly into the mixture inversion equations.

One of the great strengths of hyperspectral systems (see Chapter 11) is that direct material identification is possible (see also the discussion above). Materials that are not common in a scene but that have diagnostic absorptions distinct from the suite of end members used to model the mixture systematics can be detected and mapped in spectral mixture analysis through the use of band residual images. As discussed above, the RMS error provides a summary of the degree to which the mixture model fits the actual data. However, it does not indicate if this is distributed randomly across the wavelength region or concentrated in a few narrow absorptions. Band residuals are calculated as simply the difference between the model spectrum and the actual data, and stored as a separate file. Gillespie et al. (1990) used this approach to identify small, localized occurrences of Fe2+-bearing epidote in metavolcanic outcrops in Owens Valley. This type of analysis could be a component of the approaches discussed above for mapping the presence of specific minerals or materials that are not part of the spectral mixture model.

5.6.2 Example

The basic procedures and results of spectral mixture analysis are best illustrated by way of example (Mustard, 1994). Figure 5.12 (see color insert) is a Landsat TM scene of a well-exposed geologic terrain within the Cape Smith Fold and Thrust Belt of northern Quebec. It was constructed in the Early Proterozoic (ca. 1920–1840 my) (Parrish, 1989) during northward underthrusting of Superior Province basement (Hoffman, 1985; St.-Onge and Lucas, 1990), resulting in a large stack of normal sequence thrust faults which were reimbricated by later out-of-sequence thrusts and folded during the latest period (St.-Onge et al., 1988). The region covered by this scene includes the Archean basement (lower right corner) and fluvial to deepwater sediments (rest of the image). Many gabbro and layered peridotite–gabbro sills have intruded these sediments and the strength contrast between the sills and the sediments has resulted in the well-developed folds shown by the resistant sill units. The layered sills can be up to 500 m thick and have a consistent internal stratigraphy. A basal chilled margin of melano gabbro is overlain by pyroxene-rich peridotite that grades to peridotite and then into cyclically interlayered olivine pyroxenite and peridotite. Massive gabbro and pyroxene gabbro cap the sequence. The general chemistry and magmatic stratigraphy are very similar to those of differentiated komatitic flows nearby, and these sills are the target of mineral exploration for platinum group elements (Hynes and Francis, 1982; Barnes et al., 1992; St.-Onge and Lucas, 1990).
The TM color image in Figure 5.12 shows the well-exposed outcrops of gabbro and ultramafic sills in shades of blue, granodiorite in whitish tan, vegetated areas in reds and browns, and snow in white. The sedimentary units are not well exposed, covered largely by till, soil, and tundra vegetation. In this color composite it is possible to distinguish the outcrops readily from the vegetated regions. The results of an image-based mixture analysis are presented in Figure 5.13, with end-member spectra shown in Figure 5.14. It begins with the selection of three image end members, representing green vegetation, well-exposed gabbro from one of the sills, and shade (using an ideal shade that has a uniform reflectance of 0.01 at all wavelengths). This results in an average RMS error of 4.6 DN. The spatial distribution and abundance of these end members, shown by the fraction images (Figure 5.13a), is generally consistent with a priori knowledge, with the rock fractions concentrated on the ridge tops and vegetation on the interridge regions, and the shade image (shown in reverse stretch) approximates the illumination effects. Although the fractions are well bounded (i.e., fractions generally lie between 0 and 1), note that the RMS error image contains a significant amount of spatial information. This shows that the average RMS error of 4.6 DN is not evenly distributed among all pixels. Clearly, the regions well modeled by vegetation, gabbro, and shade (e.g., lakes, some ridges) have a low RMS error. The regions not well modeled exhibit distinct, spatially coherent patterns that can be related back to the color composite in Figure 5.12. For example, some of the well-exposed sills (light blue in color composite) are not well modeled in this solution and thus may represent rock types that have different spectral properties than the gabbro. Note the linear regions of extremely high RMS error that will appear in all subsequent results. These are snow banks and illustrate that some components, although present in the scene, can be ignored because they are not relevant to the overall objectives.

The spatial patterns of the RMS error image, coupled with the color composite, can be used to guide the selection of additional end members. A fourth end member representing tundra vegetation (NPVeg) is selected from the region of high RMS error. The mixture analysis for these four end members is shown in the second row of Figure 5.13. Comparison of the fraction images for vegetation and gabbro between the third and four end-member solutions shows that tundra vegetation was partially modeled by both the green vegetation and gabbro, while the shade image is largely unchanged. Note that the vegetation fraction image has more sharply defined distributions in the four-versus three-end-member case, and that the two vegetation fraction images together account for spectral variability due to the nonlithologic components. For this solution the average RMS error is 2.3 DN. Despite the lower average RMS error, there is still a substantial amount of spatial coherency (i.e., pixels of high RMS are grouped) and information in the error image. In this case the regions of higher error (excluding the snow banks) are concentrated in the areas of the rock outcrops, suggesting that an additional lithologic end member may be warranted.

The last row of Figure 5.13 shows the results of the mixture modeling for a five-end-member case. The fifth end member was selected from the image data in the region of highest error that is also associated with the rock outcrops as indicated by the color composite. In the resulting fraction images, note the change in the abundance patterns in the gabbro end member between the four- and five-end-member case. The distributions are much more sharply defined in the five-end-member solution, and many of the largest decreases are correlated with high-to-moderate abundances in the fraction image for the ultramafic end member. The green vegetation
Figure 5.13  Sequential mixing example. This sequence of images illustrates the basic methods of spectral mixture analysis using image end members as applied to the scene shown in Figure 5.12. The top row of images shows the fraction images derived from using three end members (spectra shown in Figure 5.14)—green vegetation, gabbro, and shade—where the shade image is displayed as (1 - shade) to better show how this end-member models primarily illumination effects. The second row of images shows the fraction images derived from using four end members: green vegetation, nonphotosynthetic vegetation (NPVeg), gabbro, and shade. The final row of images shows the results after including an end member for this area (Umaf). The average RMS error for this solution is 1.1 DN.
5.6.3 Objective Methods of End-Member Selection

In a typical approach to spectral mixture analysis, image end-member selection is achieved through an educated trial-and-error basis. An analyst has some knowledge of the field site or data set and a set of objectives for conducting the analysis. This guides the analyst to select an initial set of image end members to model the spectral image data. On the basis of these first results, the number and physical location of the image end member may be refined until an acceptable solution is achieved, as was demonstrated above (e.g., minimum error, objectives achieved). The analysis may be guided by other data processing techniques that help to understand the dimensionality of the data and inherent spectral variability. However, in many situa-
tions a more objective 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.

Repeatability can be achieved with a straightforward statistical approach such as principal components analysis or through novel applications of convex geometry (e.g., Xu and Greeley, 1992; Boardman, 1993). The latter method has been shown to have utility in determining the number of end members that are possible in a data set and to estimate the spectral properties of those end members. In this approach, illustrated conceptually in Figure 5.11d, the raw radiance or reflectance data are first transformed into ordered principal components using a minimum noise fraction (MNF) algorithm (Green et al., 1988). This determines the number of valid dimensions of the data sets. To this reduced data set, a simplex is fit to the convex hull of the n-dimensional data cloud. The faces of this simplex are regions void in one or more end members and the vertices of the simplex define the spectral properties of the end members. This method is repeatable and has distinct advantages for objective analysis of a data set to assess the general dimensionality and to define end members. The primary disadvantage of this method is that it is fundamentally a statistical approach dependent on the specific spectral variance of the scene and its components. Thus the resulting end members are mathematical constructs and may not be physically realistic.

The simplex method described above may be characterized as an outside-in approach to defining the image end members (i.e., the external surfaces of the data cloud are fit by a simplex). A technique that is the conceptual opposite has been developed by Tompkins et al. (1993, 1997) called modified spectral mixture analysis (MSMA). In this approach the end-member spectra are not prescribed but are treated as unknowns along with the fractional abundances. The model equations are the same as equations (5.10) and (5.11), but because \( F_i \) and \( R_{ib} \) are both unknowns, they must be solved nonlinearly. The nonlinear inversion employs a damped least-squares technique as presented by Tarantola and Valette (1982). In short, a starting model is provided (a suite of possible end-member spectra, estimated fractional abundances, and the image data to be modeled). Constraints on the solutions are imposed as additional equations (i.e., fractions must sum to 1.0) or as allowable deviations from the starting model (damping of solutions). Both the starting model and the constraints are based on a priori knowledge. Each successive iteration of the equations results in a calculated change in the previous model that will reduce the error of the fit. In essence the end members are “grown” from the inside of the data cloud to best fit the spectral variability, subject to the model constraints.

Both methods for objective end-member determination provide advantages over a more traditional approach. First, the correct number of end members needed to model the data effectively can be determined quantitatively. Second, the end members can be found in a manner that is objective, repeatable, and ultimately can be automated. In addition, because MSMA solutions are governed by the mixing equations and include constraints, it is both physically and statistically valid. In the absence of any a priori knowledge of the end members, this can be extremely useful in calibration. If one or more end members are already known, the overall set of solutions is more tightly constrained, so that hidden or previously unknown end members are defined more accurately. In fact, the identification of poorly represented end members is another significant strength of MSMA.
5.6.4 Linear Versus Nonlinear Models

A fundamental limitation of the linear model is that it is strictly valid only for the situation where the end members are arranged in discrete 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 small scales in intimate association with one another. When the end-member materials are intimately mixed on spatial scales smaller than the path length of photons through the medium, light typically interacts with more than one component, and the measured spectrum is a complex convolution of the end-member spectra rather than a simple linear mixture. The nonlinear effects in spectra of particulate mineral mixtures have been recognized for many years (e.g., Nash and Conel, 1974) and is an area of active research for vegetation and canopy studies (e.g., Roberts et al., 1993). A variety of methods have been developed to treat this situation, including the Hapke model for particulate surfaces (Hapke, 1981, 1993) and Camillo (1987), Pinty et al. (1990), and Verstraete et al. (1990) for plant canopies. Further discussion of nonlinear mixing in plant canopies will not be covered here. The photometric model of Hapke (1981, 1993) has been shown to be a powerful and useful model for application to nonlinear spectral mixing. The validity of the model for linearizing the mixture systematics has been demonstrated in laboratory studies of directional-hemispherical reflectance (Johnson et al., 1983, 1992) and bidirectional reflectance (Mustard and Pieters, 1987a, 1989) and shown to be accurate to approximately 5% absolute abundance. The technique has also been applied successfully to imaging spectrometer data for desert soils in Utah (Mustard and Pieters, 1987b).

The effects of nonlinear mixing on reflectance spectra can be quite dramatic, as illustrated in Figure 5.15. This is a two-dimensional plot (reflectance at 0.6 \( \mu m \) plotted against the reflectance at 0.9 \( \mu m \)) of a data cloud generated using a four-end-member mixing scenario. In Figure 5.15a, the 40 mixture points that constitute the cloud were calculated using a linear mixing model and prescribed fractions. Along the planes joining the end members (e.g., line A–C), the mixtures are in 25% increments. Visually, one recognizes this as a linear problem. The mixture points are spread along the lines joining the end members in this reflectance space at intervals proportional to their fractions. Thus the point halfway along a given line (e.g., line A–C) represents a 50:50 mixture of the end members at the vertices of the line. The systematics are very different in Figure 5.15b, where the same four end members and the 40 mixture points are shown, but here the mixture spectra were calculated using the nonlinear mixing model of Mustard and Pieters (1989), adapted from Hapke (1981). The nonlinear effects are clearly indicated by the curvilinear segments joining end members (e.g., A–C, A–D). In addition, the entire data cloud is shifted to the left against the segment C–D and toward the low-albedo end member (D). What drives this shift is the predominance of low-albedo end members in nonlinear mixing situations.

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 fractions calculated 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 (J. F. Mustard, unpublished data). 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 fit adequately with a linear model. However, the least-squares approach will minimize fitting errors using any of the end members in the equation. Thus end members not present in a mixture will be calculated to be present simply to minimize the error. In Figure 5.16 the abundances used to prepare the mixtures are shown on a ternary diagram together with calculated abundances using a linear and nonlinear mixture model. It is evident that the nonlinear model accurately predicts the modal abundances of the mixtures from the reflectance spectra. However, the linear model fractions are significantly in error, and for the enstatite-anorthite mixture a component of olivine is predicted which is not in the actual mixtures.
Results for Linear Mixing:

Olivine

○ Actual Abundances
× Calculated Abundances

Enstatite  Anorthite

Results for Nonlinear Mixing:

Olivine

○ Actual Abundances
● Calculated Abundances

Enstatite  Anorthite

Figure 5.16  Ternary diagrams of mineral abundances. The end-member mineral spectra were used to model the mixture spectra of Figure 5.10. The open circles indicate the known abundances of the prepared mixture. The solid circles represent abundances predicted from a nonlinear mixing model of the mixture spectra. The crosses indicate the predicted abundances derived from a linear mixing model of the same mixture spectra.

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 photometric properties of all end members and surfaces are required to perform the calculations. However, many common particulate materials exhibit quasi-Lambertian behavior when viewed at nadir for incidence angles up to 40°. Therefore, a Lambertian approximation should be adequate for a first-order solution. The calculation of single-scattering albedo from reflectance requires knowledge of the incidence and emergence angles of the end members and remotely acquired spectra. Spacecraft and aircraft pointing information and digital elevation models could be used to generate this information routinely. However, the most important obstacle to the application is that the particle size, composition, and alteration state of the end members are very important controlling parameters of the solutions. Neverthe-
less, through careful consideration of the nature of the remote data, and as spectral libraries become more well endowed with data, many of the obstacles may be removed.

5.7 SUMMARY AND CONCLUSIONS

Our objectives with this chapter have been to highlight spectral analysis methods and techniques that can be applied to remotely sensed data. We have concentrated primarily on major developments that have occurred since the last Manual of Remote Sensing. These developments have been primarily in the application of spectral analysis to hyperspectral imaging data. With the development of imaging spectrometers, it has become possible to utilize quantitative methods that have their origins in the analysis of high-quality and spectral resolution laboratory data. The translation of the laboratory understanding to remote applications has required adjustments in, and/or simplification of, these approaches. This is necessitated in part by the lower quality of the remotely acquired data, the complexity introduced by scaling to the larger spatial scales required for images and image pixels, and to accommodate the inherent variability of natural surfaces. Nevertheless, remote sensing systems continue to improve in quality, and it is not unreasonable to expect extremely high signal/noise ratio and radiometric fidelity from current and future systems in the near term. Thus the degree of distinction between laboratory and image analysis methods is likely to continue to diminish. With this in mind, we have attempted to link the specific approaches and techniques back to the physical fundamentals for reflectance spectroscopy that underlie them. It is our hope, therefore, that although the specific methodologies and techniques will evolve and change, the fundamental theoretical structure will endure.

We began this chapter with a short overview of some broad-scale thematic approaches to the analysis of remotely sensed data (definition and mapping of broad-scale units, identification of the presence of specific mineralogic assemblages or lithologic units, and quantification of the amount of material present). The various spectral analysis techniques presented in this chapter have different strengths and weaknesses relative to the thematic approaches. As was discussed at the beginning, before initiating any analysis strategy using remotely sensed data it is important to define the objectives of the study carefully. This can then help to identify the optimum technique or suite of techniques that will maximize the extraction of information within the available resources. Below are short summaries of the various approaches with respect to the thematic approaches.

1. Simple methods. Color composites, band ratios, and statistical transforms are all categorized as simple methods. The term simple is meant to indicate that they do not require high levels of spectroscopic knowledge and can be applied in a relatively straightforward manner. These approaches are very effective for the discrimination of broad-scale spectral units and to identify regional similarities and differences in surface spectral properties. When applied with some a priori knowledge of the surface composition, they can be correlated to specific absorption properties or spectral features in materials (e.g., ferric and hydroxyl absorptions). Simple methods achieve their maximum utility in the analysis of
2. **Feature mapping/absorption maps.** The analysis and mapping of absorption features is tied closely to the physical processes of absorption and exploits the fact that many materials exhibit unique absorptions that are diagnostic of composition. To be effective the three approaches considered (absorption parameterization, mineral mapping on the basis of absorption features, absorption band modeling) require successively higher levels of knowledge. Absorption parameterization can be automated and used to catalog the presence and properties of all absorptions in a data set, which in and of itself is useful information. The utility of this information is maximized when it can then be used as an input to more detailed analysis with reference to a spectral library, which has been similarly parameterized. The mapping of specific minerals or mineral assemblages on the basis of absorption features requires a good knowledge of the principles of reflectance spectroscopy to properly guide the selection of continuum points and links to the spectral library.

Absorption band modeling is the most flexible of the approaches under this category but requires an even greater attention to detail. Some of the approaches have been demonstrated to be effective for a wide range of materials and applications (e.g., Clark and Swayze, 1995), while others show promise but have been tested in fewer applications (e.g., Sunshine and Pieters, 1993). As has been demonstrated impressively for Cuprite, Nevada, these approaches can provide a significant enhancement for field mapping, in that specific minerals, compositions, and assemblages can be identified and mapped. These capabilities should become even greater with continued improvement in the quality of hyperspectral data from remote sensing instruments. The primary shortcomings of the methods are that they can be overly focused on absorption features and therefore may miss important features that are not part of a database or modeling experience. Also, these approaches may provide erroneous results if not applied with a concrete knowledge of the field sites and spectroscopy.

3. **Full spectral mapping.** Full spectral mapping may be characterized as spectral comparison techniques whereby the spectrum of a target material is compared against the remote observations and the results indicate where the correlations are the best. In that regard, they can be distinguished from the feature mapping techniques in that the full spectra range and suite of spectral properties are typically employed. The simpler approaches, such as binary encoding, are fast and efficient, although they are not sensitive to the subtleties of the spectral properties. In contrast, the low-probability detection algorithm can be tuned to detect subtle differences between spectra exhibited by the continuum properties. Where such differences are significant, this approach is optimal.

4. **Mixture Modeling.** When this approach is applied with proper knowledge and understanding and in field sites in which mixing is a suitable process to account for the spectral variability, abundance determinations can be estimated reliably. The approach has great utility when the surface is dominated by a few, spectrally distinct components that exhibit continuous changes in relative mixing ratios across the scene. It is also one of the premier approaches for
change detection and monitoring the composition of a surface through time. This is because spectral mixture analysis provides a consistent framework in which to analyze changes in surface abundances that can account for variability in illumination and other scene conditions. The approach is less successful for applications where a large number of discrete materials are expected (e.g., Cuprite, Nevada), due to the fact that wavelengths are highly correlated and to a first-order approximation, albedo and continuum properties dominate the least-squares mixing algorithm.

**Future Directions**

The developments in remote sensing instrumentation and spectral analysis over the last decade have been outstanding. When the first imaging spectrometers came into operation in the early 1980s, there was tremendous excitement in the remote sensing and spectroscopy communities that these new instruments would greatly improve our ability to identify and map surface compositions. However, these early systems suffered from low signal/noise ratio and other development problems, and many questioned if the incremental improvement in mapping surface compositions remotely was worth the increase in data volume, calibration requirements, and analysis effort. The community of users and instrument developers persisted, however, and as the quality of the data improved, the level of sophistication in data analysis also increased. Now there is little question that hyperspectral remote sensing systems are extremely valuable, and part of that success is due to the development of quantitative spectral analysis techniques.

Hyperspectral imaging data will become more widely used in the next decade as the current and planned instruments become operational, and this increase in data availability will open up many new and interesting applications. A key concept that we have tried to demonstrate in this chapter is that no single spectral analysis method will guarantee success. However, we have emphasized that the specific approach selected must be tied to the objectives of the analysis. If that objective is the identification and discrimination of minerals or lithologies in a complex geologic environment, an algorithm similar to that developed by Clark et al. (1990) would be recommended. If the objectives are surface abundance estimates across a terrain with relatively simple lithologic and surface units, mixture modeling would be a preferred approach. In contrast, the low-probability detection algorithms would provide the greatest opportunities for success in the identification of target materials that are poorly exposed, weakly discriminated spectrally from other components in the scene, and are not common in a given scene.

We expect that with the advent of new systems, data streams, and applications, the greater community of users will develop new and novel approaches to spectral analysis. In general, we expect that the best of these new methods will be soundly based on spectroscopic principles rather than statistical data-driven methods. The advanced algorithms and approaches covered in this chapter are still relatively new and are likely to become more refined over the next few years. The spectroscopic techniques that have emerged from laboratory investigations were developed under carefully controlled conditions but were typically limited in the number of test cases that could be examined. Now as data from remote systems approach the fidelity of
laboratory data, the opportunities to assess the utility of these methods in different environments and under a variety of environmental conditions will expand tremendously. We expect that hybrid methods will emerge and that terrain modeling (e.g., explicit incorporation of digital elevation models) and merging of data across wavelength ranges from the visible to the microwave will become more important. There is still much to learn about data exploitation and we look forward to these anticipated developments and the exciting opportunities that will emerge.


Mustard, J. F., and C. M. Pieters, 1987b. Abundance and distribution of ultramafic


Taylor, M. M., 1974. Principal components color display of ERTS imagery, in *Pro-


