Spectral Mixture Analysis of Multispectral Thermal Infrared Images

Alan R. Gillespie
Department of Geological Sciences, University of Washington, Seattle

Remote spectral measurements of light reflected or emitted from terrestrial scenes is commonly integrated over areas sufficiently large that the surface comprises more than one component. Techniques have been developed to analyze multispectral or imaging spectrometer data in terms of a wide range of mixtures of a limited number of components. Spectral mixture analysis has been used primarily for visible and near-infrared images, but it may also be applied to thermal infrared data. Two approaches are reviewed: binary mixing and a more general treatment for isothermal mixtures of a greater number of components.

INTRODUCTION

Multispectral thermal infrared images may be modeled as mixtures of a few spectral endmembers, each corresponding to a significant scene component. Spectral mixture analysis produces images of component temperatures, and images depicting the fractions of the endmember spectra needed to replicate the image. These fractions may be related to proportions of endmember materials in the scene itself through calibration to measurements made in field.

For visible and near-infrared (VNIR) images of reflected sunlight, the systematics of spectral mixture analysis have been elaborated [see Gillespie et al. (1990a) and references therein]. However, the thermal infrared (TIR) spectrum has received less attention, and mixture analysis for terrestrial targets appears to have been used only by Dozier (1981), who treated theoretically the special case in which two unisothermal endmembers of known emissivities mix to comprise the scene, and Gillespie et al. (1990b,c), who considered a more general case for linear analysis and isothermal scene elements.

Fundamental Assumptions

In conventional analyses, the scene is ordinarily taken to consist of pixel-sized elements of identifiable composition, and a meaningful emissivity spectrum is assumed to exist for each pixel. The scene elements may contain any of a vast array of constituents. If they contain multiple components, then a unique mixed spectrum is assumed to exist for each pixel. If these assumptions are correct, important components may be identified from discrete features in each spectrum.

In spectral mixture analysis, the scene is taken to consist of only a few spectrally unique components (endmembers), the number and identity of which depend upon the nature of the scene, but also the spatial scale, spectral resolution, and number of bands in the image (Adams et al., 1989; 1992).
MATHEMATICAL FRAMEWORK

If the above assumptions are valid, a set of component fractions and one or more temperatures may be deduced for each pixel. The spectrum measured for each pixel may be estimated from the spectra of the components, their temperatures and abundances. Information about unusual or unmodeled components is concentrated in "residual" images of the measured minus the estimated radiance. To date, only the simplified mixing models discussed below have been tested on terrestrial scenes.

Binary Mixing

In analyzing two-channel thermal images, it has been assumed that the scene consists of only two components, of different temperatures but known or identical emissivities. These conditions are met for several classes of remote-sensing problems, but analysis of most geological targets is infeasible because the lithologies, and hence the emissivities, are unknown and different. Dozier (1981) considered mixing for two scene components, and developed an algorithm to estimate the temperature and areal proportion of subpixel targets, using two-channel TIR data such as collected by TIROS-N (-4 μm and -11 μm). He considered the problem with and without corrections for atmospheric attenuation. In the first instance, the radiance R measured by the sensor may be described as

\[ R_i = f_i \varepsilon_i B(T_i) + (1 - f_i) \varepsilon_o B(T_o), \quad i = 1, 2, \]

where \( T \) is temperature (Kelvin), \( \varepsilon \) is the emissivity, \( f \) is the mixing fraction, and subscripts \( t, b \), and \( i \) refer to the target, background, and image channel, respectively. \( B \) is the blackbody radiance defined by the Planck equation, integrated over the image channel \( i \). \( T_i \), the apparent temperature for the pixel, must be deduced by inverting \( R_i = \varepsilon_i B(T_i) \), where \( \varepsilon_i \) is the average emissivity for the pixel.

For Eq. (1) there are two measurements: hence two unknowns may be determined. However, five variables are unknown: \( T_i, T_o, \varepsilon_i, \varepsilon_o, \) and \( f \), three of which must be assumed or found from other data or additional equations. Dozier assumed that the emissivities for target and background were the same and known (it is sufficient that they be known), leaving one variable unaccounted for. He considered two cases: (1) \( T_b \) was known; and (2) \( T_i \) and \( T_b \) were constant over multipixel image regions.

For the first case, \( T_b \) was known a priori, or found by inspection of pixels in the image for which \( f = 0 \). The pair of simultaneous nonlinear equations (1) may be solved graphically or numerically. The second case was handled by inspecting two pixels having unchanged \( T_i \) and \( T_b \), but different \( f \)'s. Rearranging Eq. (1) gives

\[ f_i = \frac{[R_i(T_i) - R_i(T_b)]}{[R_i(T_i) - R_i(T_b)]} \]

\[ i = 1, 2, \]

where \( j \) is the pixel address. The denominator is independent of \( j \), and may be eliminated by ratioing Eq. (2) for the two pixels:

\[ \frac{f_1}{f_2} = \frac{[R_i(T_{1,i}) - R_i(T_b)]}{[R_i(T_{2,i}) - R_i(T_b)]} \]

\[ i = 1, 2. \]

Thus, \( T_b \) can be computed from Eq. (3), and \( f_1 \) and \( f_2 \) from Eq. (1). A similar expression can be written for \( (1 - f) \) and \( T_b \) of Eq. (1). Smith and Rao (1972) first noted that the ratios in Eq. (3) are wavelength-invariant; hence

\[ \frac{[R_i(T_{1,i}) - R_i(T_o)]}{[R_i(T_{1,o}) - R_i(T_o)]} \]

\[ = \frac{[R_i(T_{2,i}) - R_i(T_o)]}{[R_i(T_{2,o}) - R_i(T_o)]}, \]

where \( T_e \) refers to either \( T_i \) or \( T_b \). Equation (4) has two roots, and may be solved iteratively. Once \( T_i \) and \( T_b \) are found, they may be used in Eq. (1) to find \( f \).

For case 2 (constant \( T_i \) and \( T_b \)), solution requires three instead of two adjacent pixels with the same \( T_b \) and \( T_o \), but with different mixing proportions. Analysis is based on the same ratioing tactics described above, except that in an intermediate step there are two equations—for \( (1 - f_1) / (1 - f_2) \) and for \( (1 - f_3)/(1 - f_2) \)—that are themselves ratioed to produce an expression similar to Eq. (4).

Although Dozier considered only finding temperatures and mixing proportions, for a limited range of situations, a number of important applications were encompassed. These included finding temperatures of point sources of heat, snow cover estimates, and sea-surface temperatures under a partly cloudy sky.
General Linear Isothermal Mixing

In general, mixing involves more than two endmembers, and more than two bands are required for analysis. Three assumptions must be made in analyzing these multiband TIR images in terms of general mixtures: (1) mixing is additive; (2) scene components in each pixel are isothermal; and (3) radiance \( R(\lambda, T) = f(\lambda)g(T) + c \), where \( \lambda \) is wavelength, \( T \) is temperature, \( f \) and \( g \) are functions, and \( c \) is a constant. (Planck's law, which defines radiance accurately, is not separable in terms of \( \lambda \) and \( T \).) The significance of the third assumption is that, for wavelengths \( \lambda_i \) and \( \lambda_k \) and coefficients \( a \) and \( \beta \), \( R_{ik} = aR_{ik} + \beta \), such that shape or "color" of the radiance spectrum is independent of temperature.

The first assumption is reasonable for silicate and carbonate minerals, because the TIR mean optical path length is smaller than most grains. The second assumption is unrealistic but necessary to keep the number of unknowns as small as the number of measurements. The third is met if the maximum deviation of the line \( R_{ik} = aR_{ik} + \beta \) from the locus of actual radiances is smaller than the random deviations from the locus predicted by the NEAT. For NASA's TIMS scanner (Pallucconi and Meeks, 1985), with a NEAT of > 0.1° K, the third assumption is valid for a temperature range of 300° K to > 325° K; for the future ASTER scanner to be flown on EOS, with a NEAT of 0.3° K, the assumption is valid for a range of 45° K.

Gillespie et al. (1990b,c) considered spectral mixture analysis, for an arbitrary number of image channels. Mixing was assumed to be linear, by areal proportion only, and endmember materials were required to be isothermal. The goal of analysis was to estimate proportions of endmembers, not to calculate accurate temperatures.

If the isothermal endmember materials can have different temperatures, the radiance from a scene element is

\[
R_i = \sum_{k=1}^{n} f_k e_k B(T_k) + r_i \quad \text{and} \quad \sum_{k=1}^{n} f_k = 1,
\]

where \( k \) denotes each of \( n \) endmember materials comprising the scene and \( r \) is the residual. Here, each endmember is a radiance vector \( e_k B(T_k) \).

Equation (5) differs from the corresponding equation for reflected light in that the radiance term is exponentially dependent upon temperature.

Equation (5) describes a restricted case of anisothermal mixing; in the general case, which is more realistic, for each component the temperature may vary spatially. However, the general case is of limited practical application, because the number of unknowns greatly exceeds the number of measurements. Even in Eq. (5) there are two unknowns for each endmember, \( T_i \) and \( f_i \), and measured radiances must be described as mixtures of \( \leq m/2 \) endmembers, where \( m \) is the number of image channels. Currently available scanners such as TIMS have only six channels, and this level of generality is of little practical use. Instead, Gillespie (1990b,c) reduced the level of generality by assuming isothermal pixels (the isothermal endmember materials within a particular scene element have the same temperature), rewriting Eq. (5) as

\[
DN_i = gB(T) \sum_{k=1}^{n} f_k e_k + o + r_i
\]

where the radiance values \( R_i \) have been encoded as image DN (coefficients \( g_i \) and \( o_i \)). Here, each endmember is a DN vector, \( gB(T) e + o \). The scaling coefficients are invariant with pixel address, but \( T \) is not. Measured radiances are mixtures of \( \leq m \) endmembers, allowing at least one degree of freedom each for \( T \) and \( r \). In practice, \( n \) may be estimated by the number of eigenvalues for the covariance matrix that exceed the sensor noise level, or by tentative solutions of Eq. (6) using successively greater numbers of endmembers until the root-mean-square (rms) residual \( (\Sigma r_i^2/m)^{1/2} \) is near the noise level (typically < 2 DN).

Two further simplifications were then made. In the first, the endmembers for every pixel were regarded as having the same arbitrary temperature \( T^* \), chosen to exceed the temperatures encountered in the scene. Thus the endmembers may be regarded as emissivity (\( \varepsilon \)) vectors, independent of \( T \).

In the second simplification, isothermal mixtures of the \( n \) endmember materials were approximated as mixtures of the endmember materials at \( T^* \) and an additional "virtual" cold endmember, analogous to the shade endmember in VNIR analyses (e.g., Smith et al., 1990a). The radiance from the virtual cold endmember is similar to that from a blackbody at a temperature lower than those
actually found in the scene, or from a cold surface having zero emissivity. The fraction $f$ of the virtual cold endmember accounts for the temperature information removed from the other, tangible endmembers by the first simplification. However, it should be noted that $T$ is logarithmically proportional to the complement of $f$, which varies linearly with radiance. Without the virtual cold endmember, colder, shaded slopes (for example) could be incorrectly described as mixtures of various rock and soil components, plus transpiring vegetation and/or snow.

Virtual cold is not found in the image or scene, but must be estimated from the distribution in the DN space of the data for a group of pixels comprising two or three tangible endmember materials at a range of temperatures. These data lie on mixing lines that converge towards virtual cold, although the intersection will not itself be observed and must be calculated.

In the simplified mixing model, Eq. (6) may be rewritten as

$$D_{\text{m}} = g' \sum_{k=1}^{n+1} f_k e_{\text{d}, k} + o + r_n$$  \hspace{1cm} (7)

where $e_{\text{d}, k} = 0$ (null vector) and the blackbody radiance for the tangible endmembers at $T^*$, common to all the endmembers, has been incorporated into the gain term: $g' = g D_B(T^*)$. At this point, $g'$ and $o$ have yet to be determined, although they have been estimated from engineering data. The endmembers in Eq. (7) are emissivity vectors.

It is possible at this point to calculate values of $f_k$ for each pixel of the image, given a nominal set of endmembers and the estimated values of $g'$ and $o$. However, the selection of endmembers from the image data is somewhat arbitrary. Generally, some approximation of the best set of endmembers is obtained by inspection of the image (for relatively “pure” pixels) and the cluster of data in a radiance or DN space (for evidence of mixing lines or surfaces constrained to pass through endmembers). Also, for many scenes the structure or degree of organization in the images of $f_k$ is reduced for inappropriate endmembers; therefore, some choices can be ruled out by photointerpretation. However, calculation of $f_k$ images for the image-defined endmembers is not the final step in mixture analysis.

In an image, “pure” pixels are uncommon, and most are mixed from the pure endmembers that might have meaning to the field scientist. Thus the image-defined endmembers are typically mixtures of more basic components, regardless of how well they are chosen. To make the mixture analysis meaningful, it is generally necessary to express the image in terms of these more basic components, for which emissivity spectra may be measured in the field or laboratory.

The image-defined endmembers and the reference laboratory or field emissivity spectra are related by

$$e_{\text{u}} = \frac{1}{g'} \sum_{j=1}^{n} f_j e_{\text{d}, j} - \left(o_j / g'\right) + \rho_n$$  \hspace{1cm} (8)

where subscript $k$, as before, refers to the $n$ image emissivity endmembers, the added subscript $j$, now refers to the library or reference endmembers, and $\rho$ is the residual. Equation (8) is used to select the reference spectra that fit the image endmembers best (as determined from the size of the residuals and, later, from images of $f_k$), and to determine improved estimates of $g'$ and $o$.

In all, Eq. (8) describes a set of $(mn)$ simultaneous equations and $(2m + n)$ unknowns (excluding $\rho$). For TIMS $(m = 6)$, provided $n \geq 3$, Eq. (8) is overdetermined. For actual data the matrix tends to be unstable, and the solution is sensitive to system noise; that is, there is a wide range of possible solutions for small changes in $g'$ or $o$, or for small changes in the endmember vectors (e.g., Pech et al., 1986). With iterative approximation, it is possible to reduce the ambiguity of the solution, even for noisy data (Smith et al., 1990).

In solving Eq. (8), values of $f$ are first assumed, and $g$ and $o$ are calculated, for a large number of subsets of $n$ reference spectra selected from the library (or $g$ and $o$ are assumed, and values of $f$ are calculated). Solutions yielding unreasonable values are discarded, and the mean values of the calculated variables are assumed for another round of solution. The process is repeated until few sets of endmembers are left, and successive iterations result in little change in values of $f$, $g$, and $o$. The set of spectra with the smallest residuals ($\rho$), fractions ($f$) ranging from 0 to 1, and the most meaningful identities is taken to be the set of reference endmembers.

Coefficients $g'$ and $o$ will include time-variable contributions from atmospheric attenuation and emittance, respectively. These coefficients may also vary spatially, with scan angle and in areas of
high topographic relief. It would be prudent to make model corrections for differential atmospheric effects, but this has not yet been done.

Once the reference endmembers and the calibration coefficients are known, Eq. (7) may be solved for the reference endmember fractions \( f_i \) using the reference spectra instead of the image endmembers. Fraction images, together with the residual images for each band and an overall root mean square (RMS) residual image, are the final products.

**Discussion.** Any violation of the isothermal assumption reduces the quantitative significance of \( f_i \). An example of this problem is a 1:1 areal mixture of transpiring vegetation and sand: during the day, the vegetation is colder than the land, and the preponderance of radiance comes from the sand, and the vegetation fraction appears to be quite low; however, at night the vegetation and sand may be closer in temperature, and the fractions are similar. The correct proportions are calculated only when the temperatures are the same. A second common, if less severe, example is provided by topographic shading at the subpixel scale.

A critical complication is faced in modeling TIR data that is unparalleled in earlier mixture analyses of reflected light: the shape of the blackbody spectrum changes with temperature. However, the color change is minor for commonly encountered terrestrial temperature ranges. This change may be disregarded, provided that the color aberration is less than random color errors due to instrument noise. If this is not the case (e.g., for active volcanoes with large temperature ranges), then the linear model must be modified. Ignoring color changes in the blackbody spectrum with temperature simplifies the relationship of radiance and temperature by linearizing the wavelength dependence (but not the exponential dependence of \( R \) upon \( T \)).

**EXAMPLE: ANALYSIS OF DEATH VALLEY IMAGE**

Spectral mixture analysis is illustrated using a TIMS image of Death Valley, in semiarid southeastern California. Death Valley is complex lithologically, and sparsely vegetated. It has been studied previously using multispectral thermal images (Kahle and Goetz, 1983; Kahle et al., 1984; Gillespie et al., 1984).

TIMS images with an \( \Delta T \) of 0.1–0.2° K are acquired in six bands centered at wavelengths 8.3 \( \mu \)m, 8.7 \( \mu \)m, 9.1 \( \mu \)m, 9.8 \( \mu \)m, 10.4 \( \mu \)m, and 11.3 \( \mu \)m (Palluconi and Meeks, 1985). The image of Death Valley was taken near noon, 27 August 1982. The nadir ground pixel size was \( \approx 18 \) m. The data were nominally calibrated, scan by scan, using dual on-board blackbody measurements, and geometrically resampled to rectify scan-angle foreshortening. Both the above corrections were done at the Jet Propulsion Laboratory. Differential atmospheric corrections were not made, although these could be done in the future using LOWTRAN (e.g., van den Bosch and Alley, 1990) or coregistered VNIR data (Conel, 1990). Spectral mixture analysis was done using Washington Image and Spectral Package (WISP), an interactive system developed and implemented in LISP on Symbolics computers at the University of Washington (Shippert et al., 1988).

Fraction images for scene components calculated for the Death Valley image are shown in Figure 1. These may be compared to standard enhanced color versions of the same image shown in a companion article (Gillespie, 1992, this issue). Brightness temperatures calculated from the TIMS data ranged from 319 K to 349 K, with only small portions of the mountains at lower temperatures (R. E. Alley, personal communication, 1989). Five image endmembers were employed: “virtual cold,” “vegetation,” “quartzite,” “basalt,” and “dolomite.” The “virtual cold” DN vector was \((-15, -7, -13, -9, -43, 23)\); the other endmember vectors had much larger values ranging between 150 DN and 250 DN. The name, “virtual cold,” was chosen because temperatures as low as indicated by the endmember vector do not actually occur in the image. Together, the five endmembers accounted for all the image variance not due to noise, except where there was standing water. The average RMS residual was < 1.7 DN. Calibration to reference emissivity spectra was done using field spectra taken in Death Valley (e.g., Kahle et al., 1984). The vegetation and dolomite spectra were sufficiently similar that spatial patterns in the reference-endmember fraction images were erratic, and the reference model was
rerun without dolomite, increasing the average rms residual to 3.2 DN. The four remaining fraction images are the ones shown in Figure 1a–d. Display of the fraction information in image format is an aid to photointerpretation, and the discussion below is organized from this perspective; but it should be recalled that one of the strengths of spectral mixture analysis is the numerical representation of component abundances.

The fractions in Figure 1a are related nonlinearly to temperature, and of the images shown this one most closely resembles the acquired radiance data. Important features in the scene are numbered in Figure 1a. The image spans Death Valley valley itself and includes the adjacent piedmonts (1, 2) as well as part of the Panamint range (3).

On the valley floor shallow standing water (4), calcite, halite, and silty sediments dominate. The eastern piedmont contains volcanic and gneissic gravels. The Panamints and the western piedmont are largely dolomite (5) and quartzite (6), with lesser volumes of shale (7) and volcanic rocks (8). Vegetation is generally sparse. The mountains support a pinyon-juniper woodland above creosote-bush plains. The valley floor is unvegetated.

Figure 2 shows three of the reference endmember fraction images combined as a color picture. Because the "virtual cold" fraction image was not used in Figure 2, most of the terrain information (related to temperature) is missing, resulting in the "flat" appearance compared to the
original radiance images or to Figure 1a. The color information refers to composition, not to temperature. Features referred to in the discussion below are numbered in Figure 2.

The quartzite fraction image (red component of Fig. 2) highlights alluvial fans on the Panamint piedmont that are built by streams that drain the core of the range (9). These fans appear yellow or orange in Figure 2 because they contain high fractions of quartzite and dolomite, which to TIMS spectrally resembles vegetation. Variable patterns within the quartzite-rich fans are related to age differences: The resistant quartzites are disproportionately represented on surfaces of older deposits (Gillespie et al., 1984). Adjacent fans that are dark (10) are on smaller streams that drain only the range front, which consists largely of carbonate or volcanic lithologies. Bedrock outcrops of quartzite are most evident in Figure 2 as thin yellow streaks along the Panamint range front (11); here the thin-bedded Eureka Quartzite, interstratified with carbonates, is repeated by faulting. The “vegetation” fraction (green), which encompasses all spectrally flat scene constituents, is highest on the valley floor, especially at the foot of the piedmonts (12). Like vegetation, halite and calcite approximate a graybody spectrum as viewed by TIMS. Basalt fractions (blue) are highest for the volcanics cropping out on the Panamint range front (13), and for the gneissic gravels of the eastern piedmont (14).

As noted above, temperature-related information was substantially removed from the fraction images of tangible endmember materials combined in Figure 2. Because the tangible endmember images appear to be largely independent of temperature, they probably change little in appearance with time of day or season. They thus might form a more stable basis for assessment of change—for instance, in vegetation or soil
CONCLUSIONS

Most natural scenes are mixed at the pixel scale, and it is appropriate to analyze them in those terms. Two approaches have been used: In the first, the scene was taken to consist of anisothermal binary mixtures of materials for which emissivities are known or the same; in the second, the scene was taken to consist of a larger number of isothermal endmembers of unknown emissivities.

Both approaches yield quantitative information on the proportions of materials and temperatures in a scene. Images of the mixing proportions produced by spectral mixture analysis can also be displayed as pictures for photointerpretation. These fraction images resemble those produced by decorrelation stretching. However, the fraction images have the advantage that they are thematically meaningful; that is, they depict proportions of spectrally distinct scene constituents as defined by the field scientist, not just radiant fluxes in a particular spectral band. They should also be temporally stable, provided that the isothermal assumption is met; in this they differ from simply enhanced radiance images, which vary in appearance from day to night, or even from day to day.

NASA Solid Earth Sciences Branch and EOS program provided funding. Jack Salisbury made several valuable suggestions, and Ken Watson provided an insightful and helpful review. Steve Willis at the University of Washington helped with the image processing. Assistance of the Jet Propulsion Laboratory is appreciated.

REFERENCES


van den Bosch, J. M., and Alley, R. E. (1990), Application of LOWTRAN 7 as an atmospheric correction to AVIRIS data, in Proc. 2nd Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Workshop, JPL Publ. 90–54, Jet Propulsion Laboratory, Pasadena, CA, pp. 87–95.