A Viable End-Member Selection Scheme for Spectral Unmixing of Multispectral Satellite Imagery Data

John A. Saghri*

Electrical & Computer Engineering Dept., Kuwait University, Kuwait

Andrew G. Tescher▲

Lockheed Martin M&DS, San Jose, California

Farid Jaradi and Mohammed Omran

Electrical & Computer Engineering Dept., Kuwait University, Kuwait

An important process in remote sensing is spectral unmixing which is used to obtain a set of species concentration maps known as abundance images. Linear pixel unmixing, also known as linear mixture modeling, assumes that the spectral signature of each pixel vector is the linear combination of a limited set of fundamental spectral components known as end-members. Thus end-member selection is the crucial first step in the spectral unmixing process. A conveniently parameterized method for determining the appropriate set of end-members for a given set of multispectral images is proposed. The end-members are obtained from a thematic map generated from a modified ISODATA clustering procedure that uses the spectral angle criterion, instead of the common Euclidean distance criterion. The centroids of the compact and well-populated clusters are selected as candidate end-members. The advantages of this technique over common mathematical and manual end-member selection techniques are, (1) the resulting end-members correspond to physically identifiable, and likely pure, species on the ground, (2) the residual error is relatively small, and (3) minimal human interaction time is required. The proposed spectral unmixing procedure was implemented in C and has been successfully applied to test imagery from various platforms including LANDSAT 5 MSS (79 m GSD) and NOAA's AVHRR (1.1 km GSD).

Journal of Imaging Science and Technology 44: 196–203 (2000)

Introduction

The most widely used method for extracting information on surface cover from remotely sensed imagery data is classification. Conventional classification techniques assign a single label to each pixel. The label can be any one of the known categories such as water, forest, soil, and rock. The resulting thematic map can become a very useful land cover interpretive aid provided that the imagery data is composed of pure pixels in the sense that each represents the spectral signature of only one species. A thematic map is thus appropriate for imagery data with a relatively small ground sampling distance (GSD) such as LANDSAT Thematic Mapper imagery with a 30 m GSD. For large GSD imagery, such as those from NOAA's AVHRR sensor with a 1.1 km GSD, accurate land cover estimation can only be achieved if each pixel is assigned not just to one, but several labels along with their respective concentrations in that pixel's footprint. The technique used to assign these labels and proportions is known as mixture modeling, or spectral unmixing. Spectral unmixing produces a compositional map, also known as abundance image, that provides a more complete land cover type information than a thematic map. The set of compositional maps depicts the proportions of all species present in each pixel footprint, while the thematic map identifies only the species with the highest concentration. Thus it can be argued that spectral unmixing yields a more complete classification information than a thematic map.

Recently the application of spectral unmixing has been extended to the area of multispectral image compression. Classification can provide the compression system with the additional capability to prioritize species of interest in the compression process Prioritization is achieved via minimizing the compression-induced error on the species of interest.[1] Via spectral unmixing, the multispectral image set is mathematically transformed into a set of compositional maps, known as abundance images each showing the concentration of one species in the scene. Assuming that the transformation is reversible, the actual compression can be carried out in the abundance image space rather than the original spectral image space. Prioritization is achieved via ad-

Original manuscript received August 4, 1999

▲ IS&T Member

* To whom correspondence should be addressed

justing the coding bit rate on a species by species basis depending on their significance in the final product domain. The conceptual design for a species-prioritized compression system involves,

1. applying spectral unmixing to the original image set,
2. coding the resulting prioritized species concentration maps at a relatively high bit rate,
3. coding the resulting non-prioritized species concentration maps at relatively low bit rate, and
4. at the decoder, re-mixing the decoded species concentration maps to form the species-prioritized reconstructed multispectral image set.

This concept is only feasible if the unmixing algorithm has a negligible residual error associated with it.

Spectral unmixing is commonly carried out via a linear mixture modeling approach. The basic assumption is that the signal received at the sensor from each pixel is a linear combination of the spectral contributions of all species present in that pixel's footprint. The technique allows decomposition of the scene in such a way as to recover the fractional contributions of fundamental components in the scene. This provides a means for obtaining sub-pixel information from the scene when the size of the interesting ground elements is much smaller than the image resolution. Linear pixel unmixing is much more flexible than usual "hard" classification procedures because it doesn't assign each pixel to a single class, with a consequent loss of potentially useful information. On the other hand, it produces results that are much easier to interpret than those of classical rotational transformations, such as principal component analysis. In addition to this, the assumption of linearity allows the simple mathematical treatment of the resulting abundance images so as to increase or reduce the relative importance of one or more scene components. This feature was utilized to achieve species prioritization in the bandwidth compression application.

Linear pixel unmixing requires the exact knowledge of the end-member spectra (the spectral signatures of the fundamental components). It further requires that the number of end-members be limited by the true spectral dimension of the scene (dimension of feature space). The latter is known as the condition of idebtifiability. This condition obviously limits the applicability of the linear pixel unmixing when using data from existing operational sensors. Firstly, the end-members may not necessarily correspond to physically identifiable species on the ground. Secondly, there may be more distinct species in the scene than the true spectral dimensionality of the scene. The latter constraint has recently been resolved by the method of dynamic selection of optimum end-member subset.[2] In this technique, an optimum subset of end-members is selected for spectral unmixing of each pixel vector in the scene. The subset is selected from the set of all available end-members.

The subject of end-member selection and their impact on the results of multispectral pixel unmixing has been addressed by several researchers.[2–12] In general, end-members are obtained via either mathematical methods such as principal component analysis or Gram-Schmidt orthogonalization, or manual techniques such as selecting them directly from the scene or from the library of end-member spectra. Mathematical techniques may have the advantages over manual techniques because they involve no human interaction time and, additionally, yield minimum residual errors from the decomposition process. However, the resulting end-members may exhibit negative components and/or do not correspond to spectral signatures of the physical materials present in the scene. As such, the resulting abundance images may be difficult to interpret.

In this article, a new method to obtain end-members from the scene is proposed. User-specified end-members may also imported and added to the end-members generated from the scene. The end-members are obtained from a thematic map generated from a modified ISODATA clustering procedure that uses the spectral angle criterion, instead of the common Euclidean distance criterion. The centroids of the compact and well-populated clusters are selected as candidate end-members. The advantages of this technique are, (1) the resulting end-members correspond to physically identifiable, and likely pure, species on the ground, (2) the residual error is relatively small, and (3) minimal human interaction time. The proposed spectral unmixing algorithm has been applied to test imagery from NOAA's AVHRR and Landsat MSS sensors. The performance in terms of the residual error and utility of the resulting compositional maps has been very promising.

Block Diagram of the Proposed Spectral Unmixing Process

Figure 1 shows the overall block diagram of the proposed spectral unmixing process. In the first module an unsupervised classification of the image set is carried out via a modified ISODATA clustering procedure. After reaching final convergence, the centroids of the resulting clusters are examined for their compactness and population. Compactness is measured with respect to the standard deviation of the spectral angle along each of the feature axis. The centroids of compact and well-populated clusters are selected as candidate end-members. End-members may also be externally introduced by the user and added to the list. The ensemble of internal and external end-members forms the pool of available candidate end-members for the subsequent spectral unmixing procedure. The so-called condition of identifiability limits the number of end-members to the true spectral dimension of the scene. In general, however, the number of available end-members exceeds this limit. For this reason, a strategy proposed by Maselli[2] for dynamic selection of optimum end-member subsets has been adopted. In the next module each pixel vector spectrally decomposed as a linear combination of its optimum subset of end-members. For each species represented by an end-member, the ensemble of all fractional components form a concentration map known as abundance map. The fractional concentration maps are then optimally mapped to an eight-bit integer format for display and storage purposes. The following sections describe the details of each module.

Clustering

The objective of clustering operation is to obtain the candidate end-members for the proposed spectral unmixing procedure. A modification of the well-known ISODATA (Iterative Self-Organizing Data Analysis Technique) unsupervised classification algorithm[13,14] have been adopted for this purpose.

ISODATA Classification Procedure

The ISODATA is an iterative classification method that uses the Euclidean distance as the similarity measure to cluster the data elements into different classes.[13,14] It may serve as the benchmark for all un-

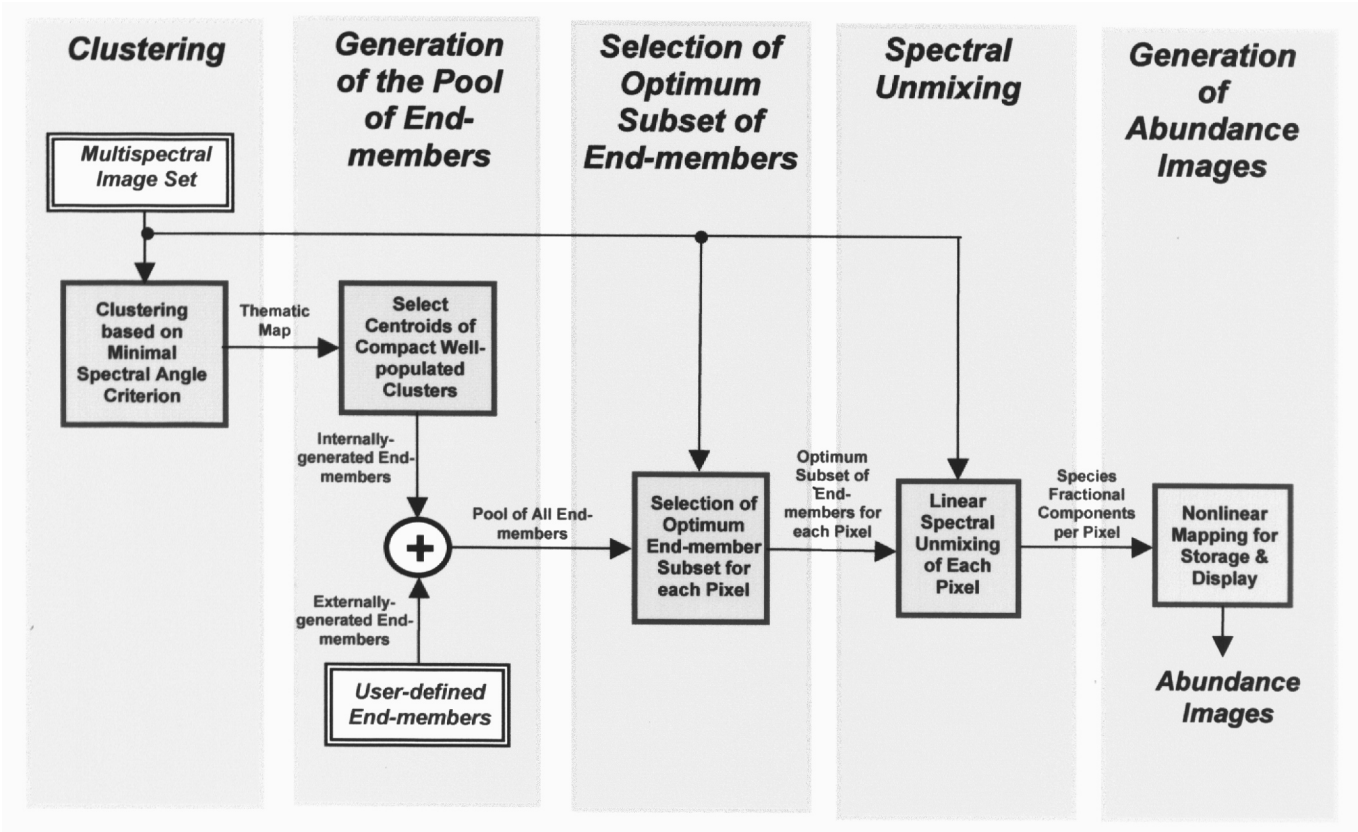

Figure 1. Block diagram of the proposed spectral unmixing process

TABLE I. General ISODATA Classification Procedure

1.	Pick K_0 arbitrary initial centroid (mean vectors)
2.	Classify the samples by assigning them to the class of the closest mean
3.	If the standard deviation of points in a cluster along a feature axis is greater than a prespecified threshold, the cluster is split in half along that axis
4.	If the Euclidean distance between the centroids of any pair of clusters is less than a prespecified threshold, the two clusters are merged into one
5.	The process is repeated with the new K_1 number of clusters until no clusters are split or merged.

supervised classification techniques. The goal of ISODATA is to divide a given set of N_p pixel vectors, i.e., spectral patterns, into exactly N_c disjoint sets, where $N_c << N_p$. By performing such a clustering, natural grouping of objects can be revealed that weren't evident before. The algorithm itself is based on the k-means algorithm, with additional heuristics that govern the splitting and/or merging of clusters.[14] The number of classes is assumed known in advance. However, no statistical information about the classes is available.

The task of the ISODATA algorithm is to perform a *labeling* operation on the set of N_p pixel vectors. Labeling refers to associating each pixel vector (feature vector) with one of the N_c clusters on the basis of the minimum Euclidean distance of the pixel vector from each cluster's center, known as *centroid*. If the number of clusters N_c is not known in advance, an iterative optimization approach outlined in Table I, known as general ISODATA procedure, may be employed to arrive at the locally optimum clustering solution.

Let N_b represent the number of spectral bands, i.e., images, in the data set. A pixel vector $\mathbf{x}_i$ can be represented in terms of its N_b components as $\mathbf{x}_i = [x_{i(1)}, x_{i(2)}, ...x_{i(N_b)}]^{tr}$. The centroid of cluster i is the mean vector $\mathbf{m}_i = [m_{i(1)}, m_{i(2)}, ...m_{i(N_b)}]^{tr}$ defined as,

$$\mathbf{m}_i = (1/n_i) \sum_{\mathbf{x} \in L_i} \mathbf{x} \tag{1}$$

where:
$\mathbf{m}_i$ = centroid of class i
n_i = number of samples in class i
L_i = subset of pixel vectors forming class i

The Euclidean distance between two pixel vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ is,

$$d(\mathbf{x}_i, \mathbf{x}_j) = \left[\sum_{k=1}^{N_b} (\mathbf{x}_{i(k)} - \mathbf{x}_{j(k)})^2 \right]^{1/2} \tag{2}$$

where

$$\mathbf{x}_i = \left[\mathrm{x}_{i(1)}, \mathrm{x}_{i(2)}, \ldots \mathrm{x}_{i(N_b)}\right]^{tr}$$

$$\mathbf{x}_j = \left[\mathrm{x}_{j(1)}, \mathrm{x}_{j(2)}, \ldots \mathrm{x}_{j(N_b)}\right]^{tr}$$

Each cluster centroid can be considered as the best representation of the ensemble of pixel vectors forming that cluster in the sense that it has the minimal sum of the squared Euclidean distances from all the cluster points. The clustering quality is thus measured in terms of the total sum of the sum of squared Euclidean distance separation for each cluster. Assume that application of the above procedure results in N_c disjoint clusters $L_1, L_2, L_3, \ldots, L_{N_c}$. The clustering quality as measured by the sum of the squared error is given by,

$$\mathrm{j}_e = \sum_{i=1}^{N_c} \sum_{\mathrm{x} \in L_i} \left[\mathrm{d}(\mathbf{x}, \mathbf{m}_i)\right]^2 \tag{3}$$

where:
- j = Total sum of the squared error
- $\mathbf{m}_i$ = centroid of class i
- d ($\mathbf{x}$, $\mathbf{m}_i$) = Euclidean distance between pixel vectors $\mathbf{x}$ and centroid $\mathbf{m}_i$
- N_c = number of samples in class i
- L_i = subset of pixel vectors forming class i

Modification of the ISODATA Clustering Algorithm

Modification to the general ISODATA clustering algorithm is accomplished via using the spectral angle measure, instead of the Euclidean distance measure, to carry out the labeling task at each iteration. That is, a pixel vector is assigned to class i if the spectral angle measured between that pixel and class i centroid is the minimum of spectral angles measured from all other class centroids. Thus the criterion to minimize the sum of the squared Euclidean distances differences in Eq. 3 is replaced by the criterion to minimize the sum of the squared spectral angle separation from the centroids.

The spectral angle α between two pixel vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ is defined as,

$$\alpha(\mathbf{x}_i, \mathbf{x}_j) = \text{arccosine} \frac{\sum_{k=1}^{N_b} (\mathbf{x}_{i(k)} \mathbf{x}_{j(k)})}{\left[\sum_{k=1}^{N_b} (\mathbf{x}_{i(k)}{}^2) \sum_{k=1}^{N_b} (\mathbf{x}_{j(k)}{}^2)\right]^{1/2}} \tag{4}$$

Note that two pixel vectors being collinear with the origin of the feature space will have zero spectral angle difference but not necessarily zero Euclidean distance separation. As such the classification result will be indifferent to multiplicative factors arising from shadow and sun angle effects.

The relationship between the spectral angle and the normalized Euclidean distance can be obtained. Let the normalized Euclidean distance between two pixel vectors $\mathbf{x}_i$ and $\mathbf{x}_j$ be denoted by d_n ($\mathbf{x}_i$, $\mathbf{x}_j$). We may then write,

$$\mathrm{d}_n(\mathbf{x}_i, \mathbf{x}_j) = \left[\sum_{k=1}^{N_b} \left([\mathbf{x}_{i(k)} / \mathrm{d}(\mathbf{x}_i, \bar{\mathrm{o}})] - [\mathbf{x}_{j(k)} / \mathrm{d}(\mathbf{x}_i, \bar{\mathrm{o}})]\right)^2\right]^{1/2} \tag{5}$$

where

$\bar{\mathrm{o}} = [0, 0, 0, \ldots, 0]^{\mathrm{tr}}$, i.e., the origin of the feature space

$\mathbf{x}_i = [\mathrm{x}_{i(1)}, \mathrm{x}_{i(2)}, \ldots \mathrm{x}_{i(N_b)}]^{\mathrm{tr}}$

$\mathbf{x}_j = [\mathrm{x}_{j(1)}, \mathrm{x}_{j(2)}, \ldots \mathrm{x}_{j(N_b)}]^{\mathrm{tr}}$

From Eqs. 4 and 5 we obtain,

$$\text{cosine } (\alpha [\mathbf{x}_i, \mathbf{x}_j]) = 1 - 0.5 [\mathrm{d}_n (\mathbf{x}_i, \mathbf{x}_j)]^2 \tag{6}$$

Equation 6 states that the minimum of spectral angle $\alpha [\mathbf{x}_i, \mathbf{x}_j]$ occurs for the minimum of the normalized Euclidean distance $\mathrm{d}_n (\mathbf{x}_i, \mathbf{x}_j)$. We conclude therefore that the result of application of the spectral angle measure in the ISODATA clustering algorithm is effectively the same as that produced via using the normalized Euclidean distance measure. In practice, the modified spectral angle-based ISODATA clustering algorithm can be implemented by applying the Euclidean distance-based ISODATA to the normalized pixel vectors.

Figure 2 shows thematic maps resulting from applying the conventional and modified ISODATA algorithms to the seven-band test imagery from Landsat™ sensor. Each map depicts a total of fourteen different classes within the same 3 km wide terrain in Montana. Inspection of the maps reveals that the map from the modified ISODATA shows a relatively larger patches of homogeneous, uncluttered, clusters and hence is simpler and more useful for interpretation of land cover types.

Obtaining Candidate End-members from the Thematic Map

The cluster centroids of the thematic map obtained above can be selected as candidate end-members for the subsequent spectral unmixing procedure. In general the larger the number of candidate end-members is the lower will be the residual error after decomposition. However, not all centroids are good candidates because they may not necessarily represent pure species in the scene. This is an important issue particularly for large footprint imagery data that is being considered for spectral unmixing. For imagery with large GSD not all pixels are pure. But all pixels, pure or otherwise, will be distributed among the final clusters in a thematic map. However the centroid of each cluster in the thematic map is the best representative of all the points in that cluster in terms of having minimum average spectral angle difference with all the points in that cluster. A smaller average spectral angle deviation means the cluster is more compact. The centroid of a compact and well-populated cluster is more likely to correspond to a pure species in the scene. Therefore, selection of the centroids as candidate end-members should be done judiciously based on the compactness and population of clusters in the thematic map. Note each of the selected end-members will have a corresponding concentration map, i.e., abundance image, following the subsequent spectral unmixing procedure.

Linear Spectral Pixel Unmixing

Linear pixel unmixing, also known as linear mixture modeling, assumes that the spectral signature of each pixel vector is the linear combination of a limited set of fundamental spectral components known as end-members. Assume that each species within a pixel footprint

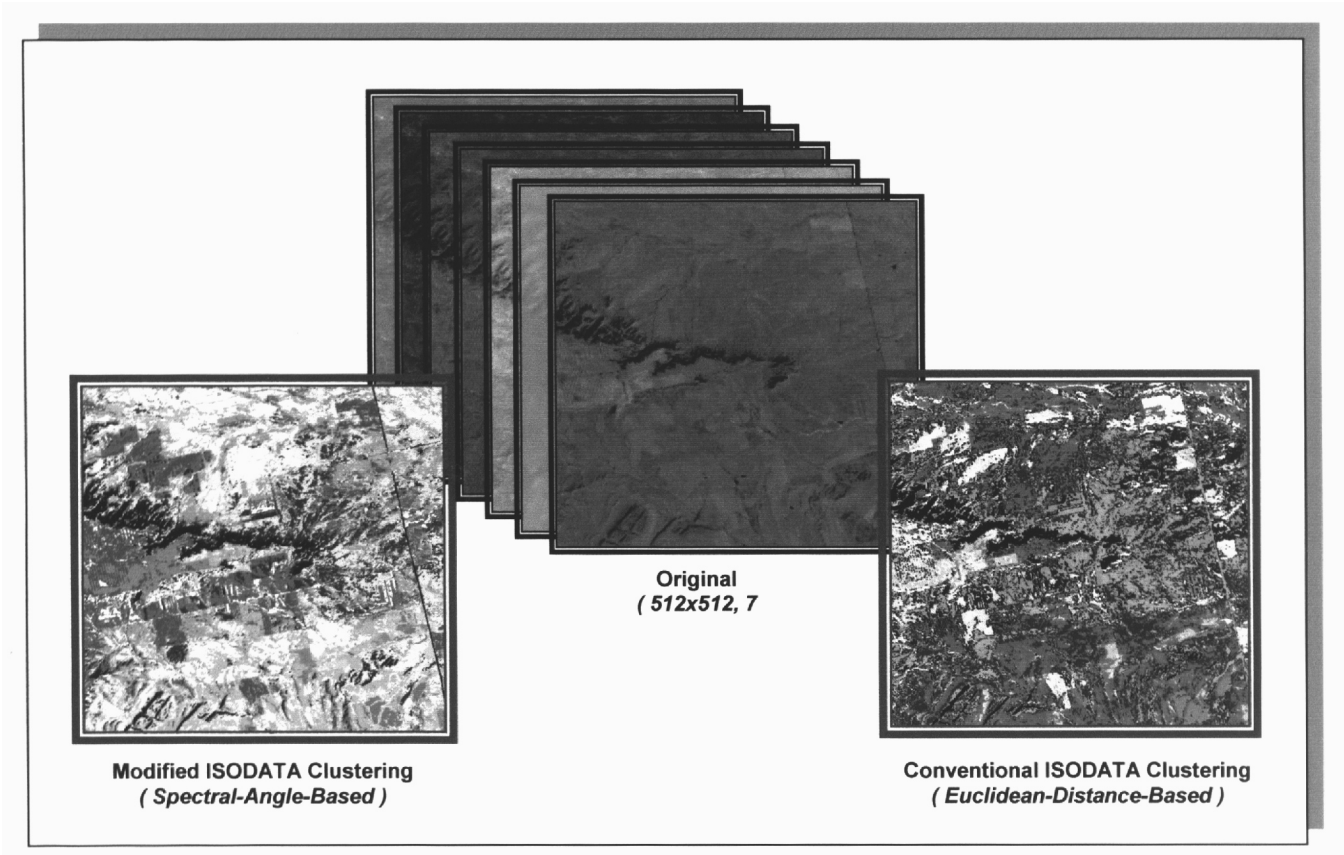

Figure 2. Fourteen-species thematic maps of Montana test image set (Size: 512 × 512 : Source: Landsat™ 7 band 8-bits Imagery).

contributes to the signal received at the satellite sensor an amount characteristic of that species and proportional to the area covered by it. The conventional spectral unmixing is modeled as,

$$\begin{aligned}\mathbf{x} &= \mathbf{M}\mathbf{f} + \mathbf{e} \\ &= f_1\mathbf{m}_1 + f_2\mathbf{m}_2 + ... + f_i\mathbf{m}_i + ... + f_n\mathbf{m}_{N_c} + \mathbf{e}\end{aligned} \tag{7}$$

where

$\mathbf{x}$	a pixel signature of N_b components
$\mathbf{M}$	$N_b \times N_c$ matrix of end-members $\mathbf{m}_{1,\ldots Nc}$
f_i	fractional component of end-member i , i.e., proportion of footprint covered by species i
$\mathbf{f}$	vector of fractional components $(f_1, f_2, \ldots f_i, \ldots f_{Nc})^{tr}$
$\mathbf{m}_i$	end-member i of N_b components
$\mathbf{e}$	residual error vector of N_b components
N_b	number of bands
N_c	number of components, $N_c \le N_d$

Provided that $N_m \le N_c$, the solution via classical least-squares estimation is,

$$\mathbf{f} = (\mathbf{M}^T\mathbf{M})^{-1}\,\mathbf{M}^T\,\mathbf{x} \tag{8}$$

Selection of the Optimum Subset of End-Members

When the number of end-members is more than the true spectral dimensionality of the scene, i.e., $N_m > N_c$, we encounter the so-called *Condition of identifiability*, which means that **f** can not be determined via Eq. 3. This situation may seriously restrict the applicability of the linear umixing operation because most operational remote sensing systems measure radiation in limited number of bands. Therefore, the scene can be decomposed into only a limited number of distinct components. For example for Landsat™ with seven spectral bands (N_b = 7), the true spectral dimension is at most five (N_c = 5) based on principal component analysis. To overcome the condition of identifiability, we adopted the method of dynamic selection of optimum end-member subset recently proposed by Maselli.[2] In this technique, an optimum subset of all available end-members is selected for spectral unmixing of each pixel vector in the scene. Thus, although not every pixel vector will have a fractional component for each end-members, the ensemble of all pixel vectors in the scene will collectively have fractional contributions for each end-member.

For each pixel vector, a unique subset of the available end-members is selected which minimizes the residual error after decomposition of that pixel vector. To determine the N_c optimum end-members for pixel vector **x**, the pixel vector is projected onto all available normalized end-members. The most efficient projection, which corresponds to the highest dot product value c_{max}, indicates the first selected end-member $\mathbf{m}_{max}$. It can be shown that this procedure is equivalent to finding the end-member with the smallest spectral angle with respect to **x**. The residual pixel signature, $\mathbf{r}_\mathbf{x} = \mathbf{x} - c_{max}$. $\mathbf{m}_{max}$ is then used to identify the second end-member by repeating the projection onto all remaining end-members. The process continues up to the identification of a prefixed maximum N_c number of end-members from the total of N_e available end-members.

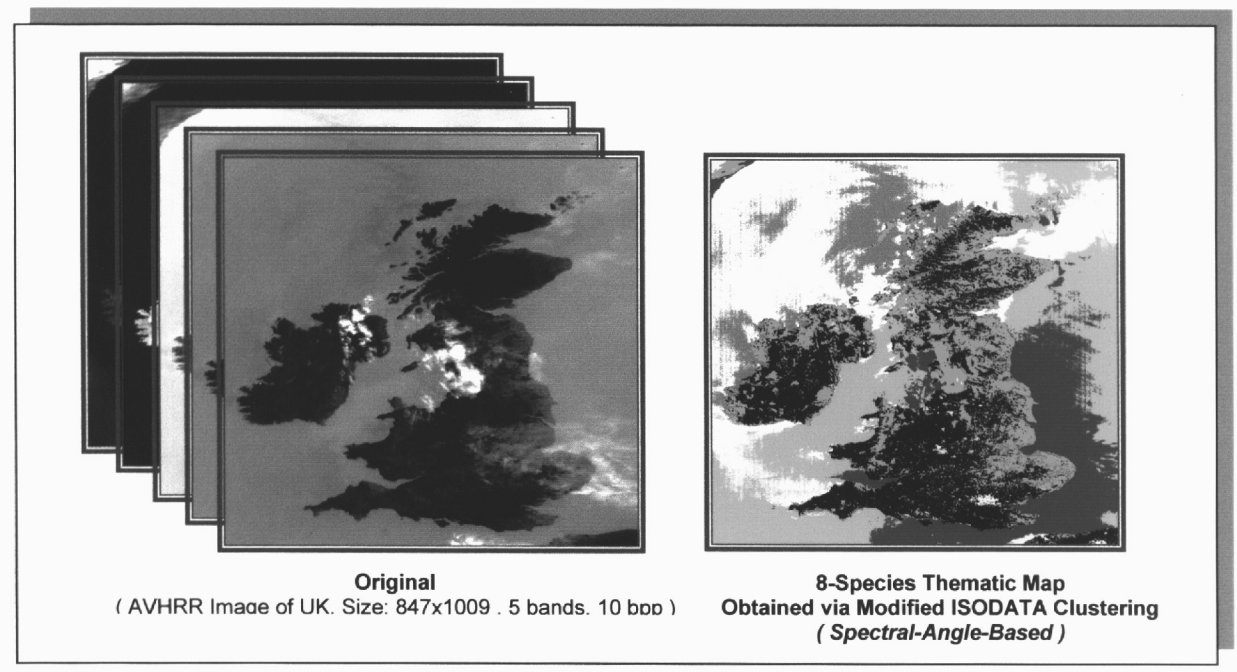

Figure 3. Test image set of UK and the corresponding thematic map (Size: 847 × 1009 : **Source:** NOAA's AVHRR 5-band 10-bits Imagery).

Displaying the Species Concentration Maps

The proposed spectral unmixing procedure produces the species concentrations as fractional, floating point, values. For display and storage purposes, these floating point maps should be converted into integer format. This requires a quantization process. For compression and archival applications, it is desirable to minimize the quantization error induced by the quantizer. For this purpose the following nonlinear mapping scheme was adopted.

$$M = \frac{255\left(f^{\exp} - (f_{\min})^{\exp}\right)}{(f_{\max})^{\exp} - (f_{\min})^{\exp}} + 0.5 \qquad (9)$$

where

M = mapped integer fractional component in the range of $0 \le M \le 255$
f = fractional component
$f_{\min}$ = minimum fractional component
$f_{\max}$ = maximum fractional component
exp = floating-point exponent parameter in the range of $0 \le \exp \le 1.0$

Note that for *exp* = 1 the Eq. 9 reduces to simple linear mapping. The optimum value of *exp* was determined, empirically, to be 0.6 for the test abundance images. The RMS quantization error for the optimum nonlinear mapping was about 30 percent lower than that for the linear mapping.

Experimental Results

The proposed spectral unmixing procedure was implemented in C and has been successfully applied to test imagery from various platforms including LANDSAT 5 MSS (79 m GSD) and NOAA's AVHRR (1.1 km GSD). In this article however, only the results for the latter is presented. The NOAA's AVHRR test imagery shown in Fig. 3 was obtained from the University of Dundee Satellite Receiving Station. It covers an almost cloud-free territory of the entire United Kingdom (UK). Each of the five spectral images (one visible, one near-infra red, and three in the thermal range) is composed of 847 × 1009 pixels. Pixels have a dynamic range of 10 bits (1024 gray levels) and correspond to a GSD of 1.1 km. The relatively large footprint of 1.1 km indicates the impurity of a large number of pixels in the scene. Hence, the test set is a good candidate for spectral unmixing to obtain sub-pixel resolution.

The pool of end-members was obtained entirely from the scene. The end-members were selected from the centroids of compact and well-populated thematic map clusters resulting from the application of the proposed modified ISODATA unsupervised clustering algorithm. The color thematic map of UK depicting 8 classes is shown in Fig. 3 along with the original image set. Figure 4 shows the set of eight abundance images resulting from the application of the discussed modified linear spectral unmixing procedure. For display purpose the fractional species concentrations were mapped to 8 bpp abundance images to obtain the depicted abundance images. The top abundance image depicts the urban areas. Figure 5 shows the resulting normalized mean and RMS residual errors using seven or eight end-members. Normalization is performed with respect to the mean (absolute values) and RMS of the test imagery data, respectively. As expected, the residual error decreases with increasing number of end-members.

To evaluate the effectiveness of the proposed technique, the same linear spectral unmixing procedure was applied to the same test data, but with different end-member selection schemes. The first experiment involved a mathematical approach. A set of five orthogonal end-members were obtained via Gram Schmidt orthogonalization procedure and used to linearly unmix each pixel vector of the scene. The second scheme involved a standard Eu-

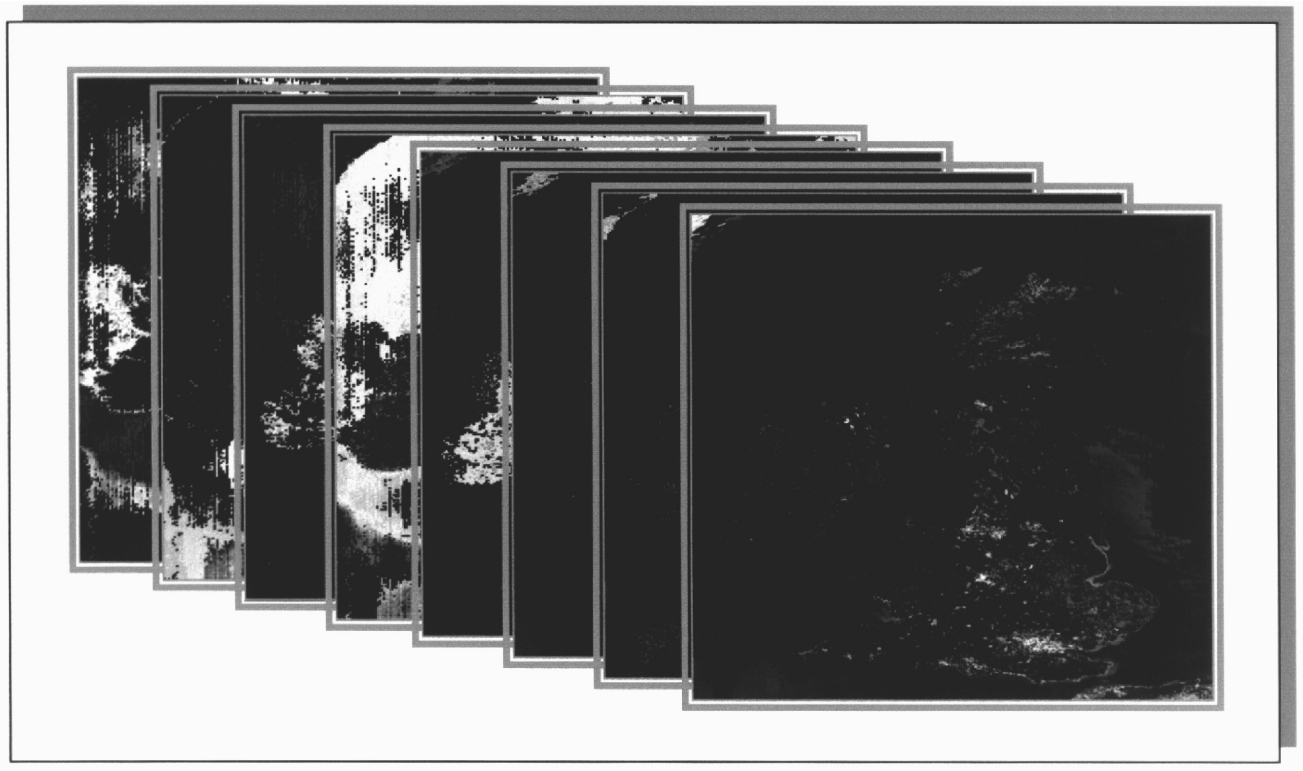

Figure 4. Species concentration maps resulting from the application of the proposed end-member selection scheme to unmix the UK test image set.

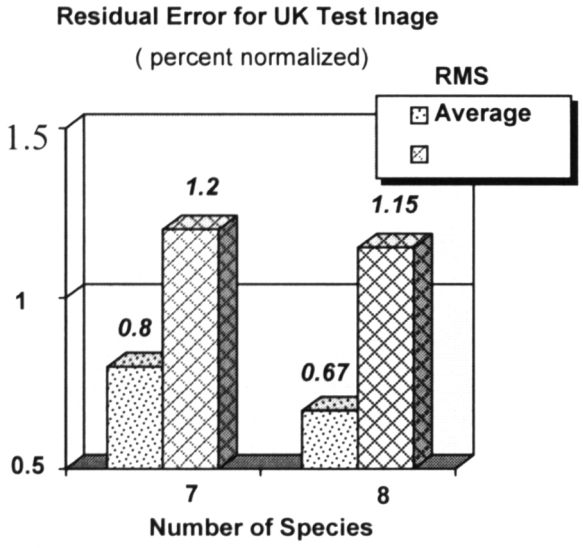

Figure 5. Residual error after decomposition for the proposed spectral unmixing scheme (using seven and eight end-members).

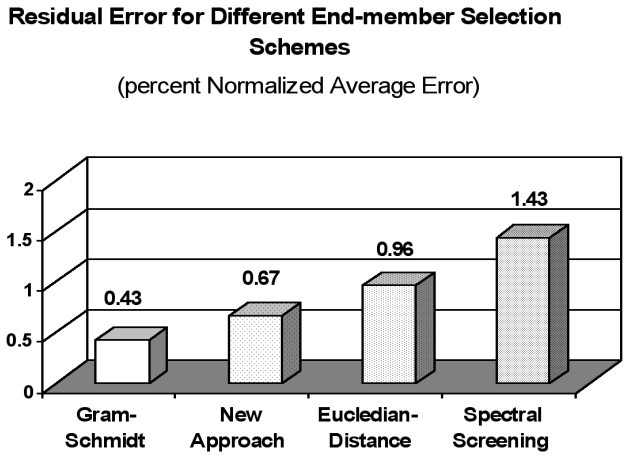

Figure 6. Residual error after decomposition for different end-member selection schemes.

clidean-distance-based method to obtain the end-members. The end-members were assigned to the centroids of the clusters formed by a standard Euclidean-distance-based clustering procedure.[13] In the third experiment the end-members were obtained using a spectral sceening procedure.[12,15] In this technique a set of pixel vectors which have unique signatures are obtained from the scene. Selection of pixel vectors is performed sequentially based on a user-defined spectral angle threshold. The pixel vectors in the image are sequentially scanned. The first pixel is considered spectrally unique and is added to the list of unique signatures. The next pixel vector in the sequence is added to the list of unique signatures if its spectral angle with respect to all the previously selected unique signatures exceeds the spectral angle threshold. Because the first unique signature can literally be any of the pixels in the scene, this technique can yield different sets of unique signatures.

Figure 6 shows the percent normalized average absolute residual error resulting from the proposed and the three traditional end-member selection schemes. As expected the mathematically optimum Gram Schmidt procedure yielded the lowest residual error. However, as discussed earlier, the end-members were not physically

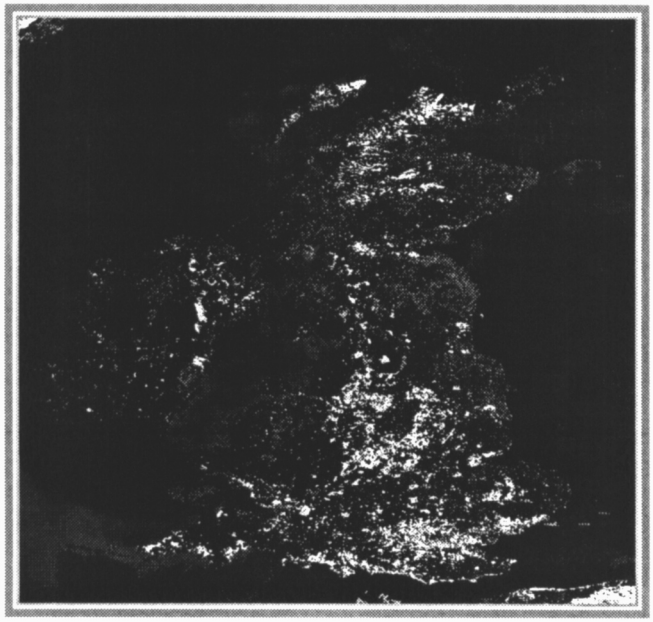

Figure 7. Urban concentration map resulting from the standard Euclidean-distance-based end-member selection scheme.

identifiable. The residual error for the spectral screening approach was almost twice as the proposed new approach. Also, unlike the proposed clustering-based approach, the end-members obtained via spectral screening are not selected based on their statistical significance, but rather their spectral uniqueness. Thus although the end-members are physically identifiable, they do not necessarily represent pure species. The Euclidean-distance-based approach also yields a higher residual error than the proposed approach. Additionally, the resulting concentration maps are not as revealing as those obtained from the proposed method. Figure 7 shows the urban concentration map of this set. Compared to the corresponding map in Fig. 4, the urban areas are not as clearly and exclusively marked in this map.

Conclusion

A conveniently parameterized method for determining the appropriate set of end-members for a given set of multispectral images was proposed. The end-members were obtained from a thematic map generated from a modified ISODATA clustering procedure that uses the spectral angle criterion, instead of the common Euclidean distance criterion. The centroids of the compact and well-populated clusters were selected as candidate end-members. The advantages of this technique over common mathematical and manual end-member selection techniques were, (1) the resulting end-members corresponded to physically identifiable, and likely pure, species on the ground, (2) the residual error was relatively small, and (3) minimal human interaction time was required. The proposed spectral unmixing procedure was implemented in C and has been successfully applied to test imagery from various platforms including LANDSAT 5 MSS (79 m GSD) and NOAA's AVHRR (1.1 km GSD).

Acknowledgement. The research reported here was sponsored in part by Kuwait University Research Administration under research grant EE 111.

References

1. J. A. Saghri and A. G. Tescher, Feature-Baesd Multispectral Bandwidth Compression, *Proc. IGARSS*, Hamburg, Germany, June 1999.
2. F. Marselli, Multiclass Spectral Decomposition of Remotely Sensed Scenes by Selective Pixel Unmixing, *IEEE Trans. Geosci. Remote Sensing* **36** [5], 1809–1819 (1998).
3. J. J. Settle and N. A. Drake, Linear Mixing and Estimation of Ground Cover Proportions, *Int. J. Remote Sensing* **14** [6], 1159–1177 (1993).
4. N. A. Quarmby, R. G. Townshend, J. J. Settle, M. Milnes, T. L. Hindle, and N. Silleos, Linear Mixture Modeling applied to AVHRR data for Crop Area Classification, *Int. J. Remote Sensing* **13** [3], 415–424 (1992).
5. A. Bateson and B. Curtiss, A Method for Manual Endmember Selection and Spectral Unmixing, *Remote Sensing of Enviornment* **55,** 229–243 (1991).
6. A. Hlavka and M. A. Spanner, Unmixing AVHRR Imagery to Access Clearcuts and Forest Regrowth on Oregon, *IEEE Trans. Geosci. Remote Sensing* **33,** 788–795 (1995).
7. Y. E. Shimabukuru and J. A. Smith, The Least Square Mixing Models to Generate Fraction Images Derived from Remote Sensing Multispectral Data, *IEEE Trans. Geosci. Remote Sensing* **29,** 16–20 (1991).
8. P. N. Bierwirth, Mineral Mapping and Vegetation Removal via Data-calibrated Pixel Unmixing, Using Multispectral Images, *Int. J. Remote Sensing*, **11,** 1999–2017 (1990).
9. J. Bowels, P. Palmadesso, M. Baumback, J. Antoniades, J. Stacka, J. Bowels, and L. J. Rickard, PHILLS: NRL's Modular Hyperspectral Imaging System, *Proc. SPIE* **2553,** 148–157 (1995).
10. P. Palmadesso, J. Antoniades, M. Baumback, J. Bowles, and L. J. Rickard, Use of Filter Vectors and Fast Convex Set Methods in Hyperspectral Analysis, *Proc. SPIE* **2553,** 1894–1912 (1995).
11. M. O. Smith, S. L. Ustin, J. B. Adams, and A. R. Gillespie, Vegetation in Deserts: A Regional measure of Abundance from Multispectral Images, *Remote Sensing of Environment*, **31,** 1–26 (1990).
12. J. Antoniades, D. Haas, P. Palmadesso, M. Baumback, and L. J. Rickard, Use of Filter Vectors in Hyperspectral Data Analysis, *Proc. SPIE* **2553,** 128–139 (1995).
13. G. Ball and D. Hall, A Clustering Technique for Summarizing Multivariate Data, *Behavioral Science* **12,** 153–155 (1967).
14. M. K. Dhodhi, J. A. Saghri, I. Ahmed, and R. Ul-Mustafa, D-ISODAT: A Distributed Algorithm for Unsupervised Classification of Remotely Sensed Data on Network of Workstations, *Journal Parallel and Distributed Computing*, *Special Issue on Software Support for Distributed Computing* **59,** 280–301 (1999).
15. J. A. Saghri, A. G. Tescher and A. Boujarwah, Spectral-signature preserving Compression of Multispectral Data, *Opt. Eng.* **38,** [12], 2081–2088 (1999).