INCORPORATING ENDMEMBER VARIABILITY INTO SPECTRAL MIXTURE ANALYSIS THROUGH ENDMEMBER BUNDLES

C. Ann **Bateson**, Gregory P. Asner, Carol A. Wessman

Cooperative Institute for Research in Environmental Sciences
University of Colorado, Boulder, CO 80302

1. INTRODUCTION

Variation in canopy structure and biochemistry induces a concomitant variation in the top-of-canopy spectral reflectance of a vegetation type. Hence, the use of a single endmember spectrum to track the fractional abundance of a given vegetation cover in a hyperspectral image may result in fractions with considerable error. One solution to the problem of endmember variability is to increase the number of endmembers used in a spectral mixture analysis of the image. For example, there could be several tree endmembers in the analysis because of differences in leaf area index (LAI) and multiple scatterings between leaves and stems. However, it is often difficult in terms of computer or human interaction time to select more than six or seven endmembers and any non-removable noise, as well as the number of uncorrelated bands in the image, limits the number of endmembers that can be discriminated. Moreover, as endmembers proliferate, their interpretation becomes increasingly difficult and often applications simply need the aerial fractions of a few land cover components which comprise most of the scene.

In order to incorporate endmember variability into spectral mixture analysis, we propose representing a landscape component type not with one endmember spectrum but with a set or bundle of spectra, each of which is feasible as the spectrum of an instance of the component (e.g., in the case of a tree component, each spectrum could reasonably be the spectral reflectance of a tree canopy). These endmember bundles can be used with nonlinear optimization algorithms to find upper and lower bounds on endmember fractions.

This approach to endmember variability naturally evolved from previous work (Bateson and Curtiss, 1996) in deriving endmembers from the data itself by fitting a triangle, tetrahedron or, more generally, a simplex to the data cloud reduced in dimension by a principal component analysis. Conceptually, endmember variability could make it difficult to find a simplex that both surrounds the data cloud and has vertices that are realistic endmember spectra with reflectances between 0 and 1.

In this paper, we create endmember bundles and bounding fraction images for an AVIRIS subscene simulated with a plant canopy radiative transfer model. The simulated subscene is spatially patterned after a subscene from the AVIRIS image acquired August, 1993 over La Copita, Texas. In addition, for comparison, we performed a traditional unmixing with image endmembers.

2. MODEL SIMULATION

A discrete ordinates plant canopy radiative transfer model was used to create AVIRIS-like image endmembers. Various aspects of the model are described in Iaquinta and Pinty (1994), Asner et al. (1997, 1998), and Asner and Wessman (1997). A brief summary description follows.

Single scattering is solved exactly while multiple scattering is simplified to a single-angle problem (zenith) (Iaquinta and Pinty, 1994). The model is designed explicitly for use with hyperspectral data, as wavelength-independent calculations (e.g. leaf angle distribution) are made only once per simulation, while those calculations requiring the leaf and stem optical properties (e.g. multiple scattering) are iterated by wavelength. The model produces top-of-canopy reflectance values from the following parameters: leaf and stem area index (LAI, SAI), leaf and stem angle

distributions (LAD, SAD), leaf and stem hemispherical reflectance and transmittance properties and soil reflectance, sun and view zenith and azimuth angles, and a hot-spot parameter for each vegetation component.

Scattering characteristics at the tissue and soil level are modeled as isotropic. Both leaf and non-photosynthetic vegetation (NPV) are included in the radiative transfer equation formulation. LAI and SAI are given on an m^2m^{-2} basis, and LAD and SAD can be modeled as erectophile, planophile, plagiophile, uniform, or as an ellipsoidal distribution with a given mean leaf angle (deWit 1965, Campbell 1986). For all analyses in this paper, solar zenith and azimuth angles were set at 20° and 0°, respectively. View zenith and azimuth angles were both set to 0°. The simulated image is based on the three dominant land-cover types found at La Copita: woody plant canopies, senescent grass canopies, and bare soil. The lack of a green grass component is realistic for La Copita in August, and has been documented in ground studies that found 95% of the graminoid biomass was dead in August 1993.

Leaf, woody stem, litter, and soil spectra gathered throughout the La Copita site were convolved to AVIRIS spectral response curves to produce 220 optical channels ranging from 400-2450 nm. LAI and mean leaf angle values gathered in August 1993 were used to constrain the range of canopy structural attributes for each simulated endmember. For each simulated AVIRIS pixel, LAI and mean leaf angle of an ellipsoidal distribution were randomly selected (Table 1). This created a range of reflectance spectra for each tree and grass endmember. A shade endmember was created based on the downwelling flux through the tree canopy (calculated in the radiative transfer model).

The tree, senescent grass, shade and bare soil endmembers were then linearly mixed by creating endmember cover fractions based on NDVI values computed for a 1993 AVIRIS subscene of La Copita. When NDVI was greater than 0.45, the tree fraction randomly varied from 70-90% cover; otherwise, it ranged from 0-69%. The shade fraction was 10% of the tree fraction. Soil and senescent grass fractions were randomly generated to sum all endmembers to 1 in each pixel. The fraction images used to direct the linear mixing of endmember spectra to produce the simulated AVIRIS image will be referred to in the sequel as the ground truth (GT) fraction images with one exception. Because the tree and shade fractions were perfectly correlated, it was not possible to derive a separate shade bundle. Hence, the image of their sum will be considered the GT tree image.

3. METHODS

3.1 Construction of Bundles

The simple scatterplot in Figure 1 illustrates the problem when pixels are mixtures of three components (Tree, Grass and Soil) with varying spectral signatures. An attempt to enclose the data in a minimal area triangle and interpret its vertices as endmembers (in this case, spectra T, G and S) will not yield spectral signatures and may in fact produce unrealistic spectra with reflectances not between 0 and 1.

What is needed is the delineation of the regions (referred to as bundles) labeled Tree, Grass and Soil in Figure 1. For the simulated La Copita scene, this delineation was done in five steps:

(1) A principal component analysis was performed on the simulated image and the data was reduced to two dimensions.

(2) Individual endmember spectra were selected automatically in the reduced dimension space using a simulated annealing algorithm. These endmembers (Figure 2) are analogous to T, G and S in Figure 1.

(3) The endmembers in (2) were significantly relocated using the manual endmember selection tool described in Bateson and Curtiss (1996) to the spectra in Figure 3, which are clearly recognizable as pure soil, tree and senescent grass. However, the data is not contained in the simplex they span.

(4) The endmembers in (3) were used as seeds to grow the soil, tree and senescent grass bundles in the

following way. Initially, each bundle contained its seed endmember. Nearby spectra were included in the bundle as long as their reflectance values ranged between 0 and 1 and they were sufficiently correlated (R=.99) to the seed endmember. Special care was taken to insure the bundles did not grow into each other and that no bundle grew into the smallest convex set containing the other bundles.

(5) In order to better contain the data within the convex hull of the three bundles, additional spectra were added to the senescent grass bundle.

The fact that realistic endmembers bounding the data cloud could not be found in steps (2) and (3) is symptomatic of endmember variability. In step (4), we limited the spectra in the tree, senescent grass and soil bundles to have reflectances bounded below by 0.02 and above by 0.5, 0.5 and 0.45 respectively. These limits were selected after examining the range of values in the image. Since the simulated image had a constant soil background, the soil bundle does not represent endmember variability but uncertainty in our knowledge of that endmember. The endmember bundles are shown in Figure 4.

3.2 Fraction Estimates

For each pixel P, we found upper and lower bounds on the tree fractional abundance by twice applying a constrained quasi-Newton algorithm to select from the bundles two sets of spectra. Each set contained one spectrum from each of the three bundles and in both cases P was contained within the triangle spanned by the spectra in the set. The first set was selected to maximize the fractional abundance of trees in P and the second set to minimize it. Repeating this procedure with soil and senescent grass, we obtained maximum and minimum images representing upper and lower bounds on endmember fractions based on the three bundles.

In addition, the simulated image was unmixed with the three image spectra having the largest fraction (~98%) of tree, senescent grass and soil in the GT fraction images. All images including the GT images are seen in Figures 6-8.

4. RESULTS

The minimum and maximum tree images (± 0.01) of Figure 8 bracket 98% of the pixels in the GT tree image. When the tree fractions computed with the image endmembers (Figure 5) were compared to the GT tree image, they had a small mean absolute error (.08). However, although the very best possible image endmembers were used to unmix the noise-free simulated scene, there were still 32 pixels with more than a 30% error in their tree fraction and 420 pixels with more than a 20% error. These large errors represent the effect that endmember variability can have on fraction estimates for individual pixels when only one spectrum is used to represent a complex landscape component like trees. In the maximum tree image, 380 of the corresponding 420 pixels differ from the GT tree image by less than 10% and in the minimum tree image 32 of the 420 pixels differ by less than 10%. Hence, the image endmember far more often underestimated than overestimated the tree fraction, but, more importantly, most of the time when large errors occurred in the tree fractions computed with the image endmembers, there was a selection of extreme endmembers from the bundles that would significantly decrease that error. Hence, the bundles approached a good characterization of endmember variability.

Soil and senescent grass minimum and maximum images (± 0.01) bracketed 94% and 99% respectively of the corresponding GT images. However, because of the low spectral contrast between soil and senescent grass, there was a close proximity between their bundles which forced their minimum images to be too small to be as useful as the Tree images.

5. CONCLUSIONS

Realistic variations in the structural attributes of trees and senescent grass together with realistic variations in tree and grass foliar optical properties are sufficient to cause errors in fractional estimates in excess of 30%.

Endmember bundles provide a conceptual framework both for understanding this variability and for quantifying the uncertainties it introduces into fractional estimates. Moreover, construction of the bundles may be accomplished without information ancillary to the AVIRIS image.

6. REFERENCES

Asner, G.P. and C.A. Wessman. 1997. Scaling PAR absorption from the leaf to landscape level in spatially heterogeneous ecosystems. Ecol. Mod. 101:145-163.

Asner, G.P., C.A. Wessman, and C.A. Bateson. 1998. Sources of variability in plant canopy hyperspectral data in a savanna ecosystem. Proc. 7th Annual JPL Airborne Earth Science Workshop.

Asner, G.P., C.A. Wessman, and J.L. Privette. 1997. Unmixing the directional reflectances of AVHRR sub-pixel landcovers. IEEE Transactions on Geoscience and Remote Sensing 35:868-878.

Bateson, C.A. and B. Curtiss, 1996. A method for manual endmember selection and spectral unmixing, Remote Sensing of Environment, 55:229-243.

deWit, C.T. 1965. Photosynthesis of leaf canopies. Agricultural Research Report 663, Pudoc Publications, Wageningen, The Netherlands.

Campbell, G.S. 1986. Extinction coefficients for radiation in plant canopies calculated using an ellipsoidal inclination angle distribution. Agricultural and Forest Meteorology 36:317-321.

Iaquinta, J. and B. Pinty. 1994. Adaptation of a bidirectional reflectance model including the hot-spot to an optically thin canopy. Proc. 6th Int'l Symp. Phys. Meas. Sign. Remote Sens., pp. 683-690.

7. ACKNOWLEDGEMENTS

This work was supported by NASA Innovative Research Grant NAGW-4689 and a grant under the NASA Earth Observing System (EOS) Interdisciplinary program (Grant NAGW-2662).

Table 1: Range of Parameter Values Used to Construct Tree and Senescent Grass Endmembers with Landscape Radiative Transfer Model.

	Range of LAI	Mean Leaf Angle of an Ellipsoidal Distribution
Tree	1.5 - 4.0	35 - 55°
Senescent Grass	1.0 - 3.0	55 - 70°

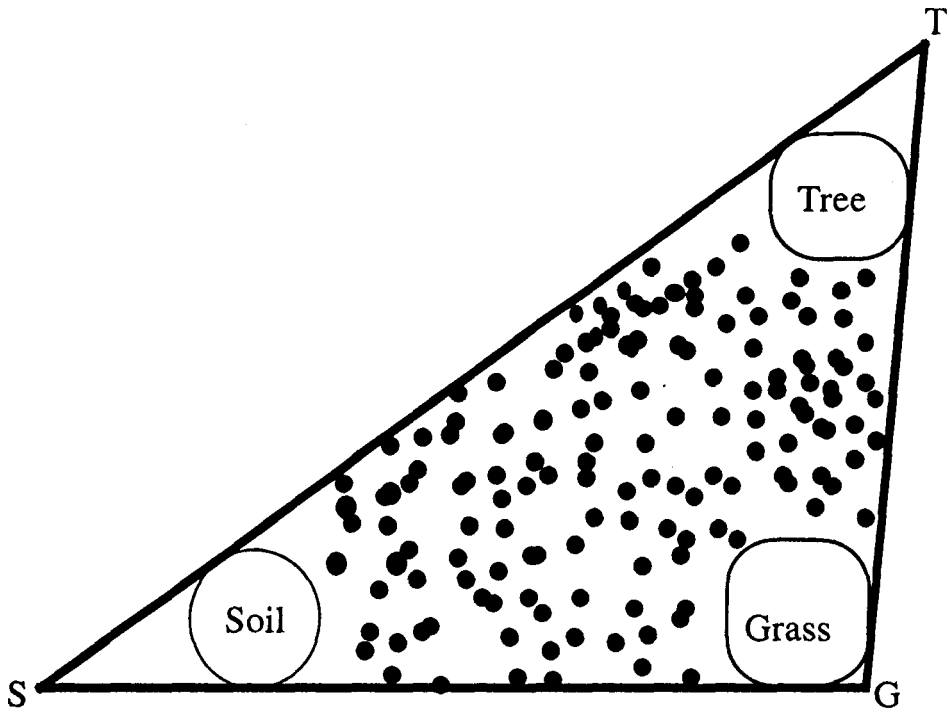


Figure 1. Scatterplot of spectral data viewed as points in the plane determined by two eigenvectors from a principal component analysis. This plot illustrates the effects of endmember variability on fitting a triangle to the data. Spectra representative of soil, trees and grass are contained within the labeled regions. The vertices, S, T and G of the smallest triangle containing the data are not within these regions and hence are not spectral endmembers. In fact, they may have unrealistic spectra with negative and superunity values.

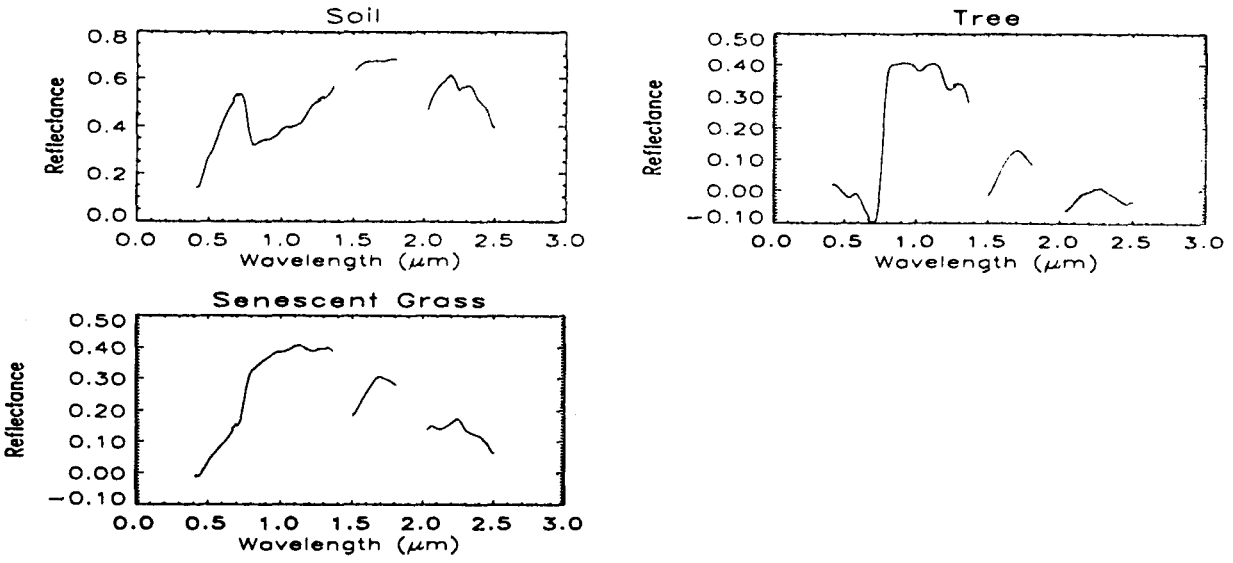


Figure 2. Endmembers automatically generated from the spectra of the simulated image. These endmembers are analogous to the vertices S, T and G in figure 1.

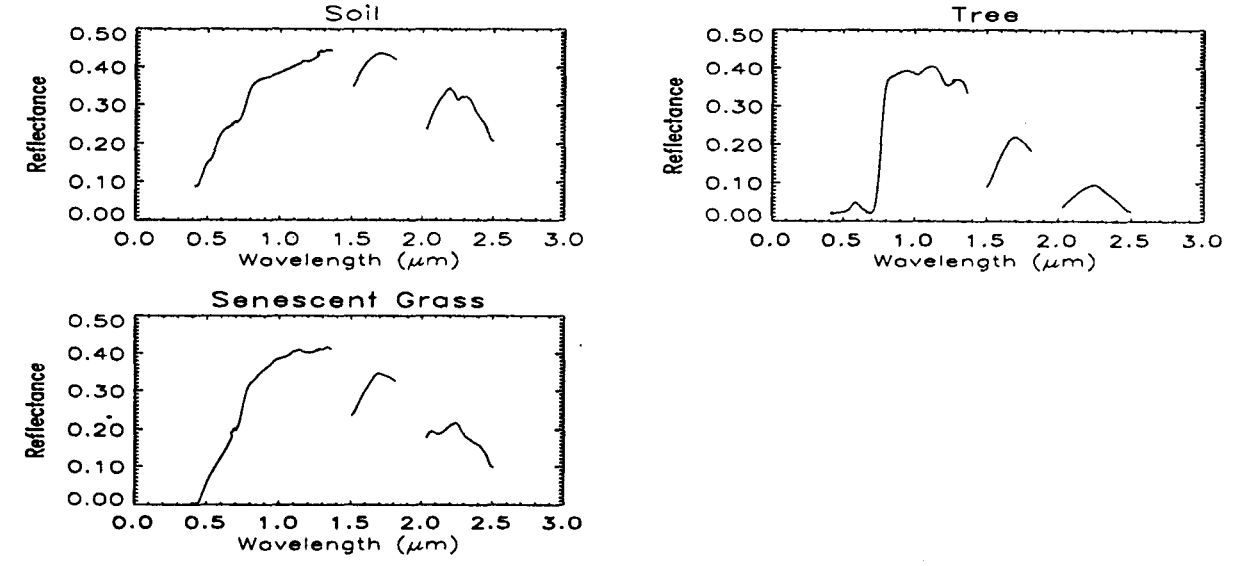


Figure 3. Endmembers selected with the manual endmember selection tool in a plane determined by two eigenvectors of a principal component analysis of the simulated image. These endmembers do not contain the data within the simplex they span.

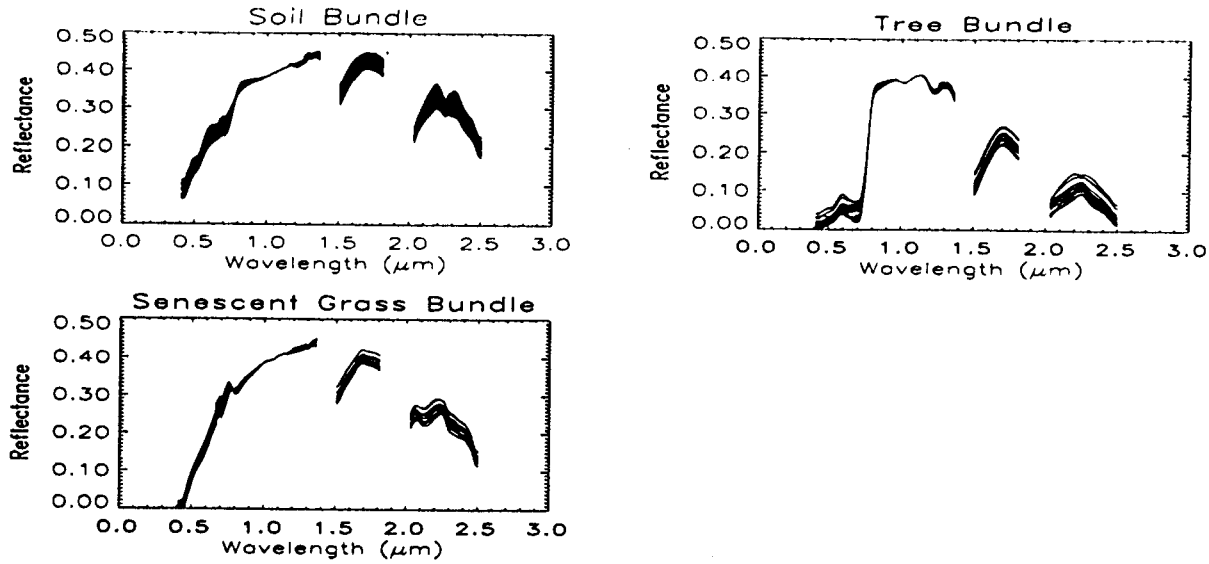


Figure 4. Endmember bundles: each bundle is grown from the corresponding endmember spectrum in Figure 3 by including nearby spectra when they are highly correlated to the seed spectrum (Pearson's correlation coefficient =.99)

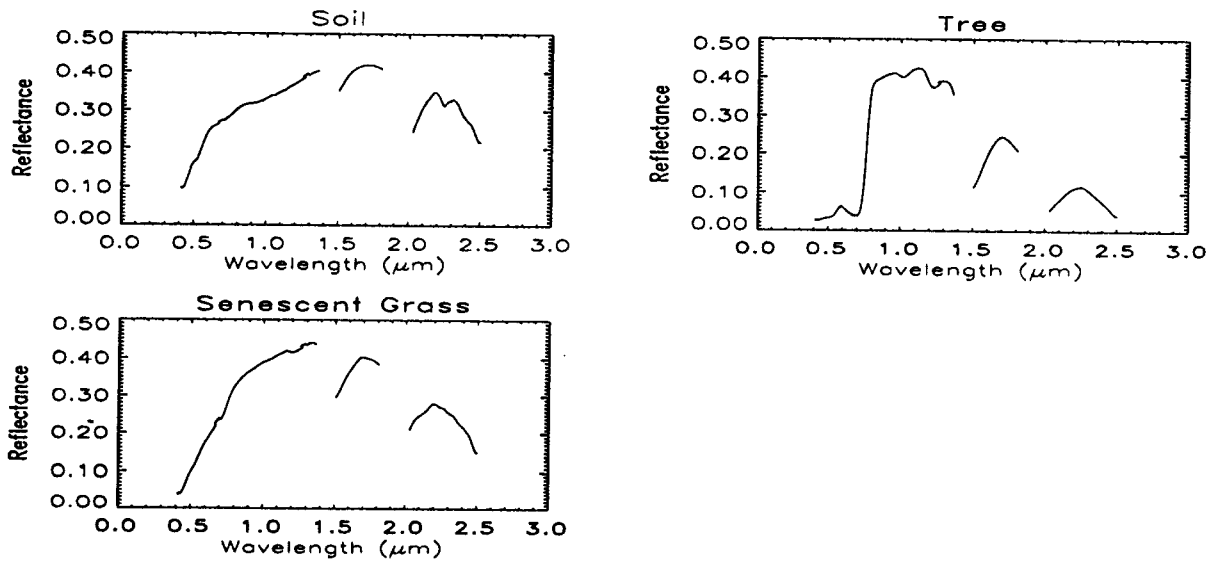


Figure 5. Endmember spectra selected directly from the image. Since these spectra are not dimension reduced through a principal component analysis, their appearance is slightly different from the spectra in figures 2-4.

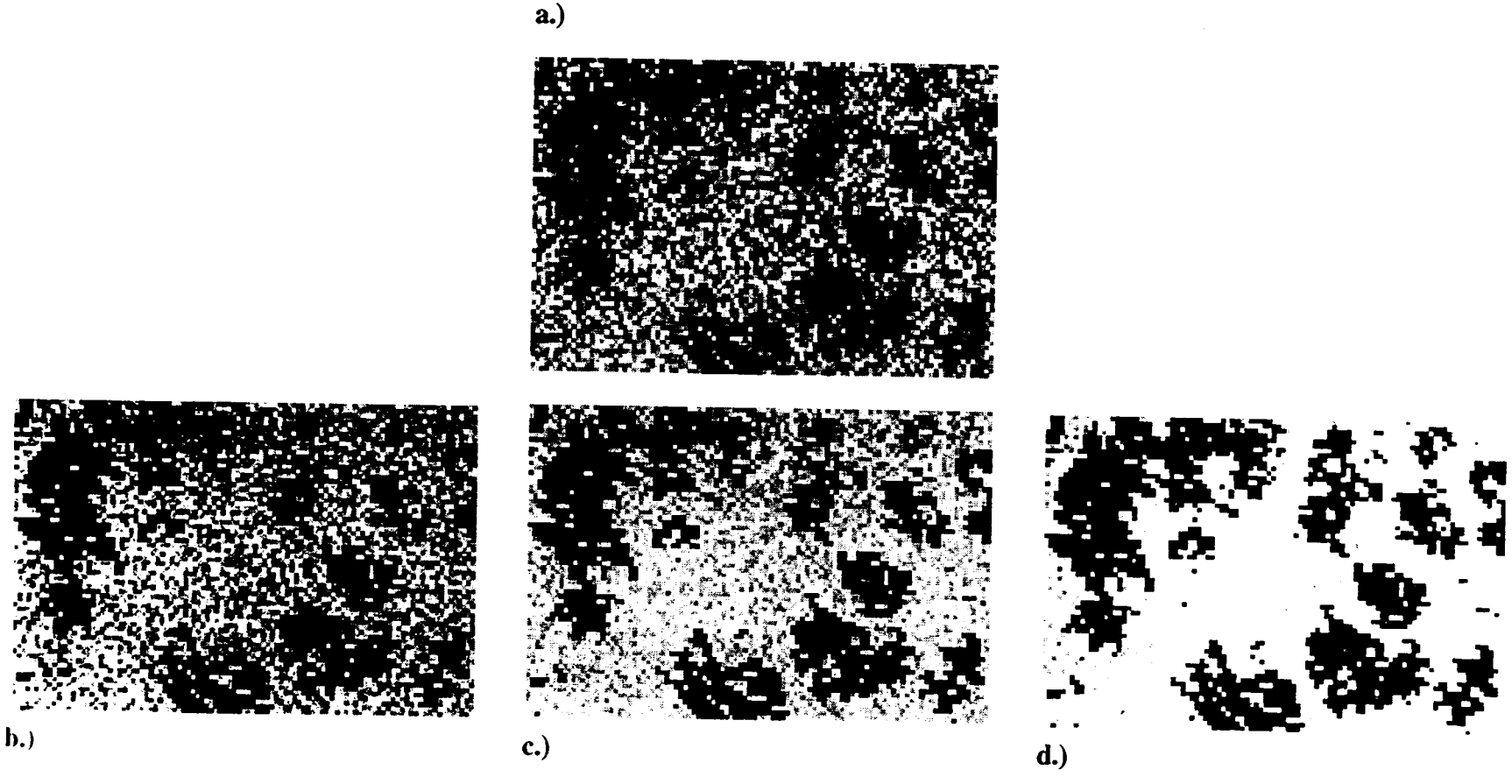
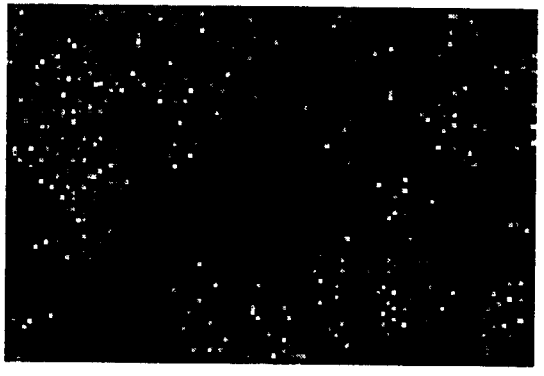
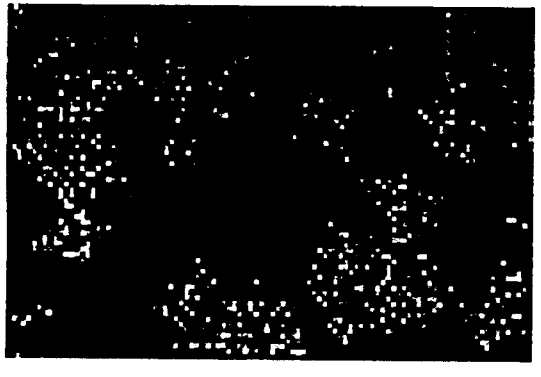


Figure 6. Tree Fraction Images: a) Fractions of the Image Tree Endmember, b.) Minimum Fractions Computed with Bundles c.) Ground Truth Fractions, and d.) maximum Fractions Computed with Bundles

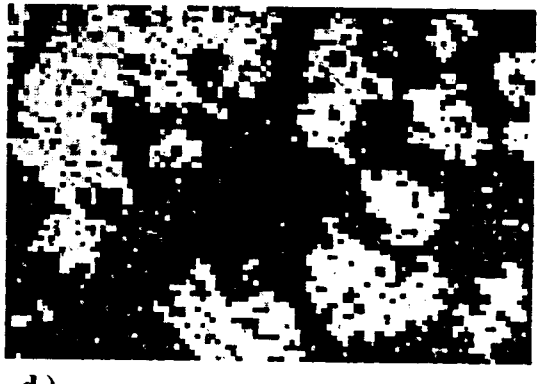
a.)





b.)

c.)



d.)

Figure 7. Soil Fraction Images: a) Fractions of the Image Soil Endmember, b.) Minimum Fractions Computed with Bundles c.) Ground Truth Fractions, and d.) Maximum Fractions Computed with Bundlles

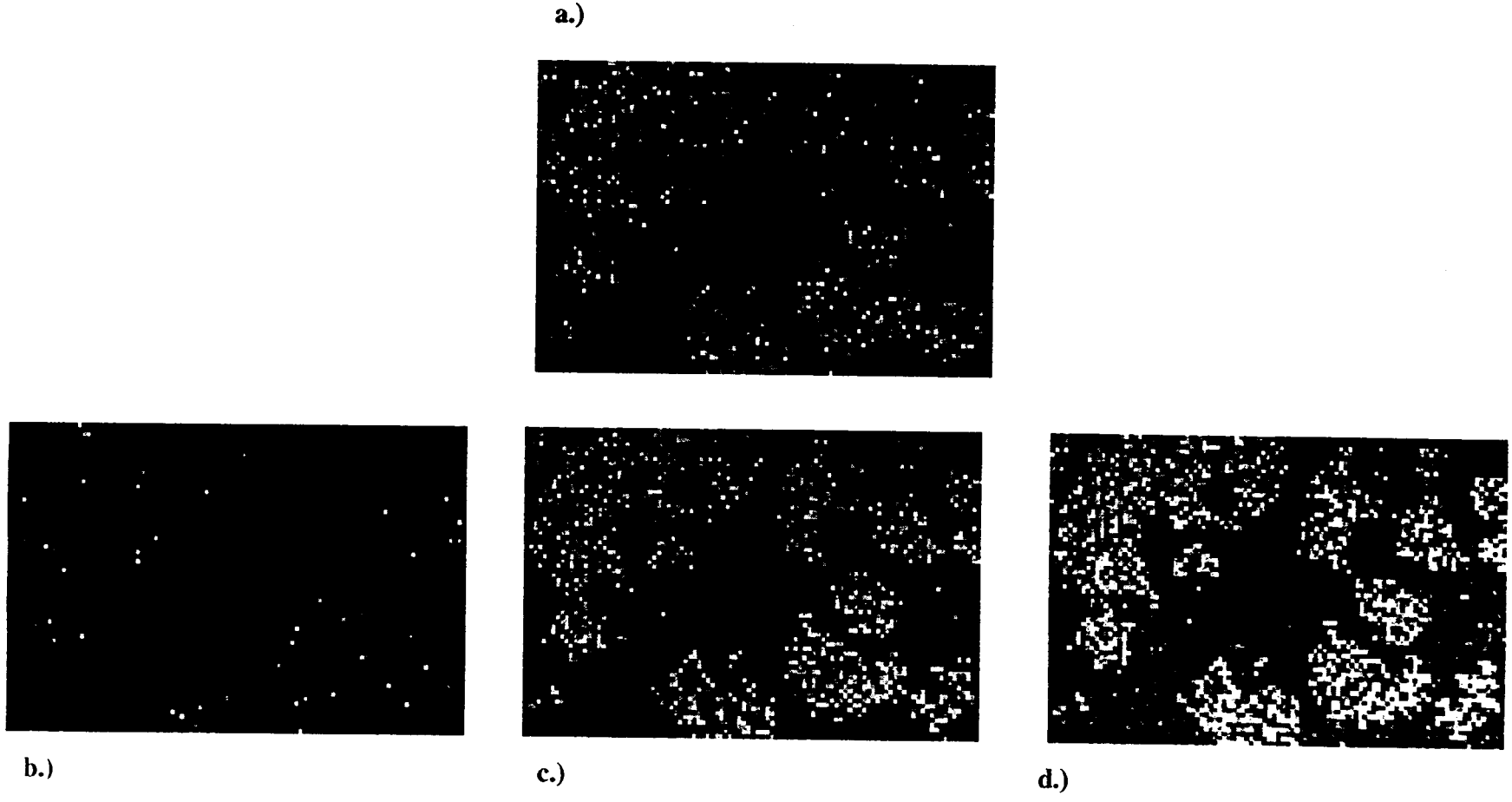


Figure 8. Senescent Grass Fraction Images: a) Fractions of the Image Senescent Grass Endmember, b.) Minimum Fraction Computed with Bundles c.) Ground Truth Fractions, and d.) Maximum Fractions Computed with Bundles