Automated Selection of Results in Hierarchical Segmentations of Remotely Sensed Hyperspectral Images

Antonio J. Plaza
Department of Computer Science
University of Extremadura
Avda. de la Universidad s/n, E-10071
Cáceres, Spain
aplaza@unex.es

James C. Tilton
Computational & Information Sciences and
Technology Office, Mail Code 606
NASA's Goddard Space Flight Center
Greenbelt, MD 20771
James.C.Tilton@nasa.gov

***Abstract*—The hierarchical image segmentation (HSEG) algorithm is a hybrid of hierarchical step-wise optimization and constrained spectral clustering. Unlike most other segmentation approaches, HSEG produces a hierarchical set of image segmentations. A single segmentation level can be selected out of the segmentation hierarchy by examining how the features or individual regions change throughout the different levels of detail. Subsequently, the selection of a single segmentation result for each region can effectively transform the segmentation hierarchy into a region-adaptive segmentation approach. The above task has previously been accomplished using supervised and time-consuming procedures. This paper presents a first step towards the automation of this process, where spatial, spectral and joint spectral/spatial features are used to investigate how regions change from one hierarchical level to the next for region identification in remotely sensed hyperspectral data sets. Comparative results are presented using Airborne Visible-Infrared Imaging Spectrometer (AVIRIS) data collected over the Salinas Valley in California.**

Keywords-Image segmentation, Hyperspectral imaging, Segmentation hierarchy, Mathematical morphology.

I. INTRODUCTION

Image segmentation is the partitioning of an image into related sections or regions. For remotely sensed images of the Earth, an example of an image segmentation would be a land-cover map that divides the image into areas covered by distinct surface covers, such as water, minerals, types of natural vegetation, agricultural crops and other types of man created development. Hyperspectral imaging is a relatively new technique in remote sensing that generates hundreds of images, corresponding to different wavelength channels, for a certain area on the surface of the Earth. For instance, the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) covers the wavelength region from 0.4–2.5 μm using 224 channels and spectral resolution of 10 nm. The incorporation of AVIRIS-type sensors on airborne/satellite platforms is currently producing a nearly continual stream of multidimensional data, and this high data volume demands efficient and unsupervised multi-channel data segmentation techniques. Specifically, in order to obtain high-quality segmentations in hyperspectral imaging, both the spatial and spectral properties of the data need to be taken into account.

The hierarchical image segmentation (HSEG) algorithm developed by Tilton [1] is one of the few available approaches in the literature that *naturally* integrates the spatial and spectral information. HSEG is a hybrid of hierarchical step-wise optimization and constrained spectral clustering that produces a segmentation hierarchy, instead of a single segmentation result. A segmentation hierarchy is a set of several image segmentations of the same image at different levels of detail in which the segmentations at coarser levels of detail can be produced from simple merges of regions at finer levels of detail. In such structure, an object of interest may be represented by multiple segments in finer levels of detail, and may be merged into a surrounding region at coarser levels of detail. A single segmentation level can be selected out of the segmentation hierarchy by analyzing the spatial and spectral characteristics of the individual regions, and also by tracking the behavior of the image segmentations throughout the different levels of detail [2]. Unfortunately, the procedure above is usually accomplished by means of supervised procedures, e.g., an analyst intensive graphical tool that allows a trained user to interactively select which segmentation resolution is most appropriate for each individual region. Although such tool can be used to label all of the various composite regions in an image, manual interaction is often subjective and extremely time consuming.

This paper represents our first step towards the automated selection of results in segmentation hierarchies. As a case study, we focus on hyperspectral image data sets. Three types of features, i.e., spatial (shape descriptors), spectral (vector-based angle metrics) and joint spectral/spatial (multi-channel morphological operations) are used to investigate how regions change from one hierarchical level to the next. Hyperspectral data sets with ancilliary information are used in experiments to evaluate the accuracy of the final segmentations, and to assess the statistical significance of regions throughout the different levels of the segmentation hierarchy.

II. HIERARCHICAL SEGMENTATION

The hierarchical image segmentation algorithm, HSEG, used in this study is unique in two major aspects. While the core of the algorithm is the relatively widely utilized hierarchical step-wise optimization (HSWO) region growing

0-7803-9050-4/05/$20.00 ©2005 IEEE.

approach [3], the HSEG algorithm uniquely allows for the merging of spatially non-adjacent regions, as controlled by the *spclust_wght* parameter. For *spclust_wght* = 0.0, HSEG is essentially the same as HSWO where only spatially adjacent are allowed to merge, for *spclust_wght* = 1.0, spatially adjacent and non-adjacent merges are given equal weight, and for values of *spclust_wght* between 0.0 and 1.0, spatially adjacent merges are favored by a factor of 1.0/*spclust_wght*.

Allowing for a range of merge priorities for spatially non-adjacent regions provides HSEG with a great deal of flexibility in tailoring the segmentation results to a particular need. HSEG also provides a selection of dissimilarity functions for determining most similar pairs of regions for merging. The currently available selection of dissimilarity functions includes functions based on vector norms, and on mean-squared error. Options for other dissimilarity functions can easily be added.

The other unique feature of HSEG is the provision of a method for selecting the most "significant" iterations from which the segmentation results are saved into an output segmentation hierarchy. The selection is performed by monitoring the behavior of the merging threshold. Whenever the ratio of the merging threshold for the current iteration divided by the merging threshold for the previous iteration exceeds a user settable threshold value, the segmentation result from the previous iteration is saved as a member of the output segmentation hierarchy. This down-selection to most significant results provides a more compact segmentation hierarchy for post-process analysis. Through this approach, HSEG provides a compact segmentation hierarchy in a single run in contrast to some other algorithms that require multiple runs to produce a segmentation hierarchy or algorithms that produce a voluminous complete segmentation hierarchy.

The allowance for the merging of spatially non-adjacent regions in HSEG leads to heavy computational demands. These demands can be significantly reduced through a recursive approximation of HSEG, called RHSEG, which recursively subdivides the imagery data into smaller sections to limit the number of regions considered at any point in the algorithm to a manageable number, usually no more than 1000 to 4000 regions. This recursive approximation also leads to a very efficient parallel implementation. The latest parallel implementation of RHSEG is so efficient that a full Landsat Thematic Mapper (TM) scene (roughly 7000 by 6500 pixels) can be processed in 5 to 10 minutes (depending on parameter settings) on a Beowulf cluster consisting of 256 2.4GHz CPUs (http://thunderhead.gsfc.nasa.gov). This is only 10 to 20 times the amount of time the Landsat TM sensor takes to collect this amount of data.

A demonstration version of RHSEG and a companion HSEGViewer program (for visualizing and manipulating the hierarchical segmentation results) is available from http://tco.gsfc.nasa.gov/RHSEG/.

III. Feature Extraction Techniques

In this section, we describe different techniques for feature extraction in the spatial and spectral domain. These features will then be used for automatic selection of features for regions at different segmentation levels. The considered approaches include spatial, spectral and joint spatial/spectral techniques.

A. Spatial Feature Extraction

Several shape analysis measurements can be used to analyze the spatial properties of regions at the individual levels of detail in the segmentation hierarchy. The considered feature measurements in this study included the *area* (number of pixels in the region), *convex_area* (number of pixels in the smallest convex polygon that can contain the region), *solidity* (proportion of the pixels in the convex hull that are also in the region, computed as *area/convex_area*) or *extent*, defined as the proportion of the pixels in the bounding box (the smallest rectangle containing the region) that are also in the region.

B. Spectral Feature Extraction

Spatial-based feature extraction does not take into account the wealth of spectral information provided by hyperspectral instruments. In order to incorporate spectral signatures into automated selection of segmentation levels of detail, we use a standard measures [4]: the spectral angle mapper (SAM). Let us consider two signatures $\boldsymbol{s}_{\mathrm{i}} = (s_{\mathrm{i}1}, s_{\mathrm{i}2}, ..., s_{\mathrm{iN}})^{\mathrm{T}}$ and $\boldsymbol{s}_{\mathrm{j}} = (s_{\mathrm{j}1}, s_{\mathrm{j}2}, ..., s_{\mathrm{jN}})^{\mathrm{T}}$, where N is the number of channels in the input data. The SAM between $\boldsymbol{s}_{\mathrm{i}}$ and $\boldsymbol{s}_{\mathrm{j}}$ is given by:

$$\mathrm{SAM}(\boldsymbol{s}_{\mathrm{i}}, \boldsymbol{s}_{\mathrm{j}}) = \cos^{-1}\left(\boldsymbol{s}_{\mathrm{i}} \cdot \boldsymbol{s}_{\mathrm{j}} / \|\boldsymbol{s}_{\mathrm{i}}\| \|\boldsymbol{s}_{\mathrm{j}}\|\right) \quad (1)$$

The SAM is invariant in the multiplication of the input vectors by constants and, consequently, is invariant to unknown multiplicative scalings that may arise due to differences in illumination and sensor observation angle, a desired feature in hyperspectral imaging. Using the SAM, we can further define a measure of spectral homogeneity within a region as follows. Let K be the number of pixels in the region R_{k}, and let $\{\boldsymbol{p}_{\mathrm{i}}\}_{\mathrm{i}=1}^{K}$ be the set of spectral signatures of the pixel vectors that compose the region. We can simply define the spectral similarity of $\{\boldsymbol{p}_{\mathrm{i}}\}_{\mathrm{i}=1}^{K}$, relative to a spectral signature $\boldsymbol{s}_{\mathrm{j}}$, as $S(R_{\mathrm{k}}, \boldsymbol{s}_{\mathrm{j}}) = (1/K)\sum_{\mathrm{i}=1}^{K} \mathrm{SAM}(\boldsymbol{s}_{\mathrm{j}}, \boldsymbol{p}_{\mathrm{i}})$. In this work, we evaluate the spectral homogeneity of each region by computing $S(R_{\mathrm{k}}, \boldsymbol{c}_{\mathrm{k}})$, where $\boldsymbol{c}_K = (1/K)\sum_{\mathrm{i}=1}^{\mathrm{K}} \boldsymbol{p}_{\mathrm{i}}$ is the centroid of R_{k}. This measure provides an indication of how similar are the spectral signatures of the pixel vectors labeled as part of the same region by HSEG. Since the algorithm may associate together pixels that are spatially disjoint but spectrally similar, the homogeneity measures above may provide better results than spatial-based metrics in subsection II-A.

C. Joint Spectral/Spatial Feature Extraction

In this subsection, a combined spectral/spatial approach for feature extraction is described. The approach is based on mathematical morphology, an image processing technique with two basic operations: erosion and dilation. These operations are respectively based on the replacement of a pixel by the neighbor with the maximum and minimum digital value, where

the pixel neighborhood is given by a so-called structuring element (SE). In order to extend the operations above to hyperspectral images, we impose an ordering relation in the set of pixel vectors lying within an SE, designed by B, by defining a cumulative distance between one particular pixel $\boldsymbol{f}(x, y)$ and all the pixel vectors in the spatial neighborhood given by B (B-neighborhood) as follows:

$$D_B[\boldsymbol{f}(x,y)] = \sum_i \sum_j \mathrm{SAM}[\boldsymbol{f}(x,y), \boldsymbol{f}(i,j)] \quad (2)$$

where (i, j) refers to spatial coordinates in the B-neighborhood. Based on the distance above, the extended erosion of $\boldsymbol{f}$ by B selects the B-neighborhood pixel vector that produces the minimum value for D_B:

$$(\boldsymbol{f} \Theta B)(x,y) = \left\{\boldsymbol{f}(x+i', y+j'), (i', j') = \arg\min_{(i,j)} \{D_B[\boldsymbol{f}(x+i, y+j)]\}\right\} \quad (3)$$

where the argmin operator selects the pixel vector is most highly similar, spectrally, to all the other pixels in the B-neighborhood. On other hand, the extended dilation of $\boldsymbol{f}$ by B selects the B-neighborhood pixel vector that produces the maximum value for D_B:

$$(\boldsymbol{f} \oplus B)(x,y) = \left\{\boldsymbol{f}(x-i', y-j'), (i', j') = \arg\max_{(i,j)} \{D_B[\boldsymbol{f}(x+i, y+j)]\}\right\} \quad (4)$$

where the argmax operator selects the pixel vector that is most spectrally distinct to all the other pixels in the B-neighborhood. Based on the above operations, we define a measure of spectral/spatial homogeneity at a given pixel [4] as follows: $\mathrm{MEI}(x,y) = \mathrm{SAM}[(\boldsymbol{f} \oplus B)(x,y), (\boldsymbol{f} \Theta B)(x,y)]$. In this work, we use the mean of MEI scores of the pixels in a region R_k as a measure of its spectral/spatial homogeneity.

IV. Experimental Results

A. Data Description

The hyperspectral scene selected for experiments is a portion of a 2001 AVIRIS data set taken over an agricultural test site located in Salinas Valley, California. The scene consists of 512 lines by 217 samples, with 154 spectral bands after removing the water absorption and noisy bands. The data include vegetables, bare soils and vineyard fields with sub-categories as given in Fig. 1, which shows the entire scene and a sub-scene of the dataset. The subscene, called "Salinas A" and outlined by a rectangle in Fig. 1, comprises 83x86 pixels and is dominated by directional classes. Figures of the ground-truth take at the time of the data acquisition are also displayed. One of the most interesting features of the Salinas data set is that it represents a hyperspectral analysis scenario dominated by directional classes with very similar spatial and spectral properties. For instance, the romaine lettuce is at different weeks since planting and with growth increasingly covering the soil, which results in slightly distinct spectral signatures. This is a challenging segmentation scenario (in particular, for an unsupervised segmentation approach). In order to facilitate our exploration of the segmentation hierarchy based on spatial and spectral features, we focus on the analysis of the lettuce_romaine fields present in the "Salinas A" subscene.

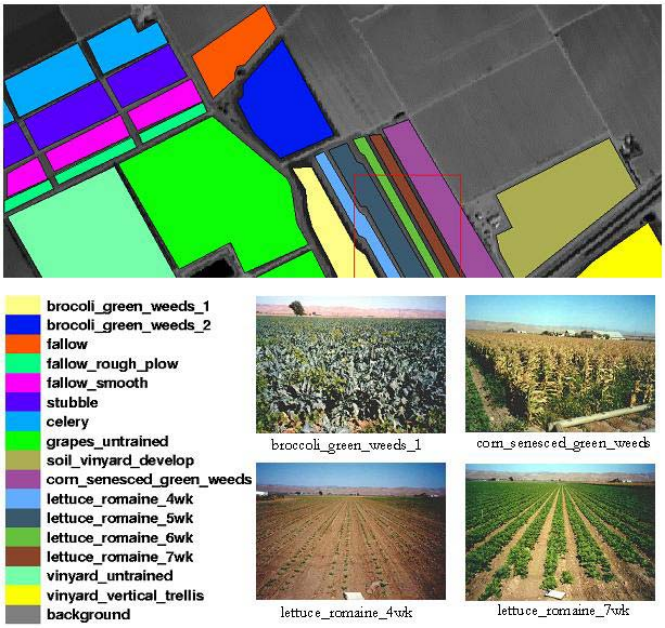

Figure 1. AVIRIS data set collected over Salinas Valley in California.

B. Analysis of the AVIRIS Salinas Data Set

In order to carry out a preliminary analysis of segmentation accuracy, we first ran HSEG on the "Salinas A" data set and used ground-truth information in Fig. 1 to compute the true and false positive rates for each region and segmentation level produced by HSEG (see Table I).

TABLE I. True (TPR) and False (FPR) Positive Rates for Several Regions at Different Segmentation Hierarchy Levels

Region	Level 1		Level 2		Level 3		Level 4	
	TPR	FPR	TPR	FPR	TPR	FPR	TPR	FPR
lettuce_4wk	0.25	0.00	0.43	0.01	0.75	0.03	0.85	0.17
lettuce_5wk	0.25	0.01	0.43	0.02	0.64	0.03	0.99	0.11
lettuce_6wk	0.12	0.00	0.33	0.02	0.99	0.05	0.99	0.15
lettuce_7wk	0.26	0.00	0.62	0.01	0.95	0.05	0.95	0.12

As shown by Fig. 2, most regions evolved from an instance of under-segmentation to levels where over-segmentations and false positives were clearly visible. Therefore, an automated selection of the best segmentation level for each region is highly desirable. Also, the scores in Table I demonstrate that level 3 may exhibit the level of segmentation detail that better fits available ground-truth information. An important question at this point is: what kind of features should be extracted from image objects in order to automatically select a single segmentation level out of the segmentation hierarchy?

In the following, we explore different techniques to extract features able to describe spectral and spatial properties of objects in remotely sensed hyperspectral data. In order to evaluate if shape measurements can provide useful information for the selection of regions at different levels, we computed the metrics in section III-A for all spatially connected regions in Fig. 2, and found that only the *solidity* parameter was able to

provide an indication about the compactness of each region at the different segmentation levels. In all cases, regions at segmentation levels 3 and 4 showed the highest compactness scores. However, it is clear from ground-truth information in Fig. 1 that, out of the extracted connected components at level 4, only one corresponds to an optimal segmentation level for the region, while the other ones can be considered as a false positive detections. The false positives at segmentation level 4 also showed high compactness scores. Subsequently, the *solidity* alone cannot be used as a measure to select a single segmentation result out of the hierarchy.

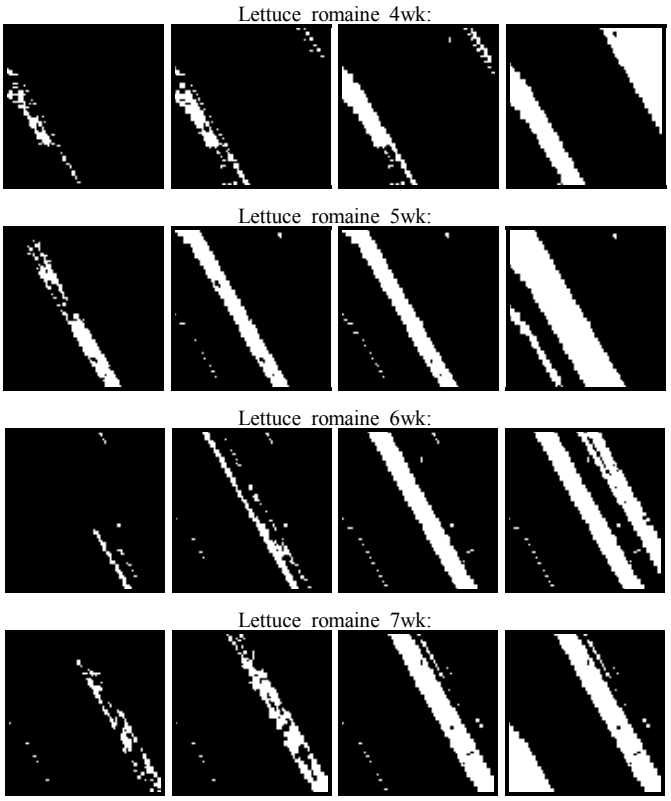

Figure 2. Tracking of lettuce_romaine regions from level 1 to level 4 (left to right) in the segmentation hierarchy produced by HSEG for "Salinas A."

TABLE II. SPECTRAL HOMOGENEITY SCORES FOR REGIONS AT SEGMENTATION LEVELS: 3 AND 4 PRODUCED BY HSEG

Level	lettuce_4wk	lettuce_4wk	lettuce_4wk	lettuce_4wk
3	0.92	0.91	0.93	0.92
4	0.56	0.69	0.71	0.58

In order to resolve the issues above, we resort to the spectral information contained in the original hyperspectral image. Since the HSEG algorithm clustered together pixels that are spatially disjoint, here we consider each region not as a set of spatially connected components as in the previous experiment, but as a set of spectrally similar pixel vectors in the multi-dimensional space comprised by the input data. Table II shows the scores produced by the spectral homogeneity metrics in section III-B for the regions at segmentation levels 3 and 4. As shown by Table II, the spectral homogeneity scores for the regions in segmentation level 3 are close to optimal, but false positive regions at segmentation level 4 resulted in significantly lower spectral homogeneity scores. Interestingly, spectral homogeneity metrics allowed us to automatically discard false positive HSEG detections based on spectral properties of the data alone.

The approaches explored thus far consider spatial and spectral information separately. To conclude this section, we investigate a joint feature selection approach that relies on simultaneous exploitation of spatial and spectral information. Table III shows the morphological eccentricity scores (defined in section III-C) for different segmentation hierarchy levels. A disk-shaped (isotropic) structuring element was considered, where the radius of the disk was set to the width in pixels of each field, computed using ground-truth information.

TABLE III. JOINT SPECTRAL/SPATIAL HOMOGENEITY SCORES.

Region	Level 1	Level 2	Level 3	Level 4
lettuce_4wk	0.48	0.54	0.99	0.65
lettuce_5wk	0.61	0.94	0.94	0.61
lettuce_6wk	0.47	0.63	0.96	0.81
lettuce_7wk	0.44	0.71	0.98	0.63

Scores in Table III provide a measure of spectral/spatial consistency that exploits both the spatial properties (through the width of the structuring element) and spectral information (spectral homogeneity of the classes). A general requirement of multidimensional morphological operations, however, is to carefully set the spatial properties of the structuring element in order to obtain the desired performance. A method for automated selection of an optimal structuring element at each pixel was recently developed in [4] in order to alleviate the above constraint in general-purpose hyperspectral applications.

V. CONCLUDING REMARKS

Unlike most other segmentation approaches, the HSEG segmentation approach produces a hierarchical set of image segmentations. The potential of segmentation hierarchies remains largely unexplored in many application areas such as remotely sensed hyperspectral imaging, which can greatly benefit from automated techniques able to exploit segmentation hierarchies in a region-adaptive fashion. In this paper, several feature extraction techniques in the spatial and spectral domain have been proposed in order to investigate how regions change from one level to another in a segmentation hierarchy. Our experimental results provided several intriguing findings that may help data analysts in selection of feature extraction approaches for automating the exploitation of segmentation hierarchies in specific applications.

REFERENCES

[1] J. C. Tilton, "Image segmentation by region growing and spectral clustering with a natural convergence criterion," *Proc. IGARSS*, Seattle, WA, pp. 1766-1768, July 6-10, 1998.

[2] J. C. Tilton, "Analysis of hierarchically related image segmentations," *Proc. IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, Greenbelt, MD, pp. 60-69, Oct 27-28, 2003.

[3] J. M. Beaulieu and M. Goldberg, "Hierarchy in picture segmentation: A stepwise optimal approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 2, pp. 150-163, 1989.

[4] A. Plaza, P. Martínez, J. Plaza and R. Pérez, "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 466-479, 2005.