https://ntrs.nasa.gov/search.jsp?R=20060013023 2019-08-29T21:36:53+00:00Z

Source of Acquisition
NASA Goddard Space Flight Center

Utilizing Hierarchical Segmentation to Generate Water and Snow Masks to Facilitate Monitoring Change with Remotely Sensed Image Data

James C. Tilton[1*], William T. Lawrence[2], and Antonio J. Plaza[3]

[1]Computational & Information Sciences and Technology Office, Mail Code 606.3, NASA Goddard Space Flight Center, Greenbelt, MD 20771 James.C.Tilton@nasa.gov

[2]Department of Natural Sciences, Bowie State University, Bowie, MD 20715 wlawrence@bowiestate.edu

[3]Department of Computer Science, University of Extremadura, Avda. De la Universidad s/n, E-10071, Cáceres, Spain aplaza@unex.es

*Corresponding Author

Utilizing Hierarchical Segmentation to Generate Water and Snow Masks to Facilitate Monitoring Change with Remotely Sensed Image Data

ABSTRACT

The hierarchical segmentation (HSEG) algorithm is a hybrid of hierarchical step-wise optimization and constrained spectral clustering that produces a hierarchical set of image segmentations. This segmentation hierarchy organizes image data in a manner that makes the image's information content more accessible for analysis by enabling region-based analysis. This paper discusses data analysis with HSEG and describes several measures of region characteristics that may be useful analyzing segmentation hierarchies for various applications. Segmentation hierarchy analysis for generating land/water and snow/ice masks from MODIS (Moderate Resolution Imaging Spectroradiometer) data was demonstrated and compared with the corresponding MODIS standard products. The masks based on HSEG segmentation hierarchies compare very favorably to the MODIS standard products. Further, the HSEG based land/water mask was specifically tailored to the MODIS data and the HSEG snow/ice mask did not require the setting of a critical threshold as required in the production of the corresponding MODIS standard product.

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 is a labeled map that divides the image into areas covered by distinct earth surface covers such as water, snow, types of natural vegetation, types of rock formations, types of agricultural crops and other man created development. In unsupervised image segmentation, the labeled map may consist of generic labels such as region 1, region 2, etc., which may be converted to meaningful labels by post-segmentation analysis.

Utilizing image segmentation as a first step in an image analysis scheme makes possible a more reliable labeling of image objects based on region feature values, which may now include region shape and texture. This is an appropriate approach for the analysis of natural image objects, since these objects are usually composed of varied patches and not discrete spatial elements like image pixels.

A segmentation hierarchy is a set of several 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. This is useful for applications that require different levels of image segmentation detail depending on the particular image objects analyzed. In such a structure, an object of interest may be represented by multiple image segments in finer levels of detail in the segmentation hierarchy, and may be merged into a surrounding region at coarser levels of detail in the segmentation hierarchy. If the segmentation hierarchy has sufficient resolution, the object of interest will be represented as a single region segment at some intermediate level of segmentation detail. The segmentation hierarchy may be analyzed to identify the hierarchical level at which the object of interest is represented by a single region segment. The object

may then be identified through its spectral and spatial characteristics. Additional clues for object identification may be obtained from the behavior of the image segmentations at the hierarchical segmentation levels above and below the level at which the object of interest is represented by a single region.

Segmentation hierarchies may be formed through a region growing approach to image segmentation. In region growing, spatially adjacent regions iteratively merge through a specified merge selection process. Hierarchical Step-Wise Optimization (HSWO) is a form of region growing segmentation in which the iterations consist of finding the best segmentation with one region less than the current segmentation (Tilton and Cox 1983, and Beaulieu and Goldberg 1989). The best segmentation is defined through a mathematical criterion such as a minimum vector norm or minimum mean squared error. An augmentation of HSWO, called HSEG (for Hierarchical Segmentation), was introduced by Tilton (1998) in which spatially non-adjacent regions are allowed to merge controlled by a threshold based on previous merges of spatially adjacent regions. HSEG also includes a method for selecting the most "significant" iterations from which the segmentation result is saved to form an output segmentation hierarchy.

Until recently, successful analysis of HSEG generated segmentation hierarchies has been accomplished only through the use of an analyst intensive graphical tool that allows a trained user to visualize and interact with the segmentation hierarchies. While it is possible to interactively label an image using such a tool, the manual interaction is subjective and time consuming. This paper describes some initial attempts towards the automated labeling of objects based on their representation in HSEG generated segmentation hierarchies. As a case study, we focused on twelve MODIS data sets with a spatial resolution of 1.0 km^2 centered approximately on the Salton Sea in southern California. The data sets were from January 31,

2003 through February 28, 2005 and included broad seasonal changes in vegetation canopy characteristics, including those due to drought and fire-induced change. While our eventual goal is to develop techniques for monitoring land-use and land- cover change in this and similar time series data sets, this paper concentrates on the use of segmentation hierarchies for generating land/water and snow/ice masks from MODIS data, and compares and contrasts these results to the corresponding MODIS standard products.

Hierarchical Segmentation

The hierarchical image segmentation algorithm that was used in this study (HSEG) is based upon the relatively widely utilized hierarchical step-wise optimization (HSWO) region growing approach of Beaulieu and Goldberg (1989), which can be summarized as follows.

- Initialize the segmentation by assigning each image pixel a region label. If a pre-segmentation is provided, label each image pixel according to the pre-segmentation. Otherwise, label each image pixel as a separate region.
- Calculate the dissimilarity criterion value between all pairs of spatially adjacent regions, find the pair of spatially adjacent regions with the smallest dissimilarity criterion value, and merge that pair of regions.
- Stop if no more merges are required. Otherwise, return to step 2.

HSEG differs from HSWO in two major aspects. First, the HSEG algorithm allows for the merging of spatially non-adjacent regions, as controlled by the *spclust_wght* parameter. For *spclust_wght* = 0.0, only spatially adjacent regions are allowed to merge, as in HSWO. However, for *spclust_wght* > 0.0, HSEG allows merges between spatially non-adjacent regions. For *spclust_wght* = 1.0, merges between spatially adjacent and non-adjacent regions are given equal weight. For values of *spclust_wght* between 0.0 and 1.0, merges

between spatially adjacent regions are favored by a factor of 1.0/*spclust_wght*. Allowing for a range of merge priorities for spatially non-adjacent regions uniquely provides HSEG with a great deal of flexibility in tailoring the segmentation results to a particular need.

Another way HSEG differs from HSWO is in the provision of an approach for selecting the most "significant" iterations from which the segmentation results are saved into an output segmentation hierarchy. In this approach, the behavior of a certain criterion is monitored (either the maximum merging threshold, a global region dissimilarity function, or the global region standard deviation). When the ratio of the criterion value for the current iteration divided by the criterion value 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 (e.g., the segmentation algorithm provided with the popular eCognition package, see http://www.definiens-imaging.com/) or algorithms that produce a voluminous complete segmentation (HSWO).

HSEG also provides a selection of dissimilarity functions for determining most similar pairs of regions for merging. The available selection of dissimilarity functions are based on vector norms, and on mean-squared error. Options for additional dissimilarity functions can easily be added.

The merging of spatially non-adjacent regions in HSEG leads to heavy computational demands. These demands are significantly reduced through a recursive approximation of HSEG, called RHSEG, which recursively subdivides the imagery data into smaller

processing windows (sections) to limit to a manageable number the number of regions considered at any point in the algorithm (usually in the range of 1000 to 4000 regions). RHSEG includes a provision to blend the results from the subsections to avoid the production of segmentation artifacts along the processing window seams. This recursive approximation also leads to a very efficient parallel implementation. This parallel implementation of RHSEG is so efficient that full Landsat Thematic Mapper (TM) scenes (approximately 7000 by 6500 pixels) can be processed in 2 – 8 minutes 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 that the Landsat TM sensor takes to collect this amount of data.

A demonstration version of RHSEG along with a companion HSEGViewer program (for visualizing and manipulating the HSEG or RHSEG segmentation hierarchy results) is available from http://tco.gsfc.nasa.gov/RHSEG/. The package includes a detailed user's manual with a more detailed description of HSEG and RHSEG along with a complete discussion of program parameters. The user's manual also includes a tutorial describing the use of HSEGViewer program.

Data Sets Analyzed

We obtained MODIS Terra MOD021KM calibrated radiance data and MOD03 geolocation data for 12 MODIS granules, as listed in Table 2. Initially we acquired only bands 1-7, which represent the primary bands for land studies. Seasonal data sets were acquired from early 2003 through early 2005. The acquisitions were meant to capture both seasonal change and evidence of large chaparral fires that occurred in Southern California in fall 2003. We bounded our search with the geographic coordinate spatial search, using 32°N, 30°N, 117°W and 116°W as our limits. Data sets were selected based on seasonality and

cloud cover criteria over the Southern California region (Figure 1). The seven reflective band MODIS data sets were converted from swath to georectified grid, from HDF to binary BSQ, uniformly subset so that all 12 data sets were coincident at a 1.0 km^2 spatial resolution, and centered approximately on the Salton Sea in southern California. Additionally, we obtained a more complete set of data for two MODIS granules (Id., A2003301.1825 and A2004080.1825). For these we obtained all 36 spectral bands (reflective and emissive) along with the cloud mask data (MOD35) and the land/sea mask from the geolocation data (MOD03).

METHODS

In this study, RHSEG was run using *spclust_wght* = 0.1, eight nearest neighbor connectivity, and the Global Maximum Merge Threshold (GMMT) convergence criterion. Dissimilarity criteria based on vector norms, image mean squared error, and a SAR (synthetic aperture radar) speckle noise criterion were available in the versions of HSEG and RHSEG used in this study. In this study, we used the Square Root of Band Sum Mean Squared Error dissimilarity criterion, $d^{1/2}_{BSMSE}(X_i, X_j)$. For image subsets or segments X_i and X_j

$$d^{1/2}_{BSMSE}(X_i, X_j) = \left[\frac{n_i n_j}{(n_i + n_j)} \sum_{b=1}^{B} (\mu_{ib} - \mu_{jb})^2 \right]^{1/2}, \quad (1)$$

where n_i (n_j) is the number of pixels in image segment X_i (X_j), u_{ib} (u_{jb}) is the mean value for the b^{th} spectral band of image segment X_i (X_j), and B is the number of spectral bands. We chose this criterion since it favors the merging of smaller regions into larger ones. The derivation of this dissimilarity criterion is provided in Appendix A. Program defaults were used for other parameters. The acronyms and notation used throughout this paper are summarized in Table 1.

Data Analysis with RHSEG

For most applications, RHSEG produces a hierarchical segmentation set that must be analyzed in a post-processing step to select a particular segmentation out of the hierarchy. This segmentation can either be from one level of the segmentation hierarchy or a blending of segmentations for several levels of the segmentation hierarchy. Either type of segmentation can be produced using the HSEGViewer program mentioned earlier. Using the HSEGViewer program an analyst exploits visual cues, external knowledge about the scene, and intuition to select an appropriate segmentation out of the segmentation hierarchy. An automatic, non-interactive approach requires the use of quantitative features to determine the appropriate segmentation. Features explored in this study are described in the following section.

Features for Post-Analysis of Segmentation Hierarchies

Region Features

A simple region feature that was used in this study was the Maximum Merge Threshold (MMT_i) encountered by the RHSEG region growing process for merges involving region i up until the current iteration. A large value for this feature indicates that the region is relatively inhomogeneous.

The Band Maximum Standard Deviation ($BM\sigma_i$) of region (here, region i) was used in this study as an indication of the variability of a region at a particular hierarchical segmentation level. For regions consisting of two or more pixels, the standard deviation for spectral band b, σ_{ib}, is:

$$\sigma_{ib} = \sqrt{\frac{1}{n_i - 1} \sum_{x_p \in X_i} (\chi_{pb} - \mu_{ib})^2} = \sqrt{\frac{1}{n_i - 1} \left[\sum_{x_p \in X_i} (\chi_{pb})^2 - n_i (\mu_{ib})^2 \right]} \quad (2)$$

where n_i is the number of pixels in the region, x_p is a pixel vector (in this case, a pixel vector in region X_i), χ_{pb} is the image data value for the b^{th} spectral band of pixel vector x_p, and μ_{ib} is the region mean for the b^{th} spectral band of region i:

$$\mu_{ib} = \frac{1}{n_i} \sum_{x_p \in \chi_i} \chi_{pb}.$$

The band maximum standard deviation for region i is then defined as:

$$\mathrm{BM}\sigma_i = \max\{\sigma_{ib} : b = 1,2,\ldots,B\} \quad (3)$$

where B is the number of spectral bands.

Region dissimilarity criteria can be used to indicate the character of a region at a specific level in the segmentation hierarchy. In particular, a region dissimilarity criterion that was used in this study is the square root of the Band Sum Mean Squared Error, $\mathrm{D}_{\mathrm{BSMSE}}^{1/2}(X_i)$. For region X_i containing n_i pixels this is:

$$\mathrm{D}_{\mathrm{BSMSE}}^{1/2}(X_i) = \left[\frac{1}{(n_i - 1)} \sum_{x_p \in X_i} \sum_{b=1}^{B} (\chi_{pb} - \mu_{ib})^2 \right]^{1/2}. \quad (4)$$

The above region dissimilarity criterion is calculated versus the original image data values (χ_{pb}). As an alternative, this criterion can be calculated versus another region, such as the region at the first (0^{th}) level of the segmentation hierarchy:

$$\mathrm{D}_{\mathrm{BSMSE}}^{1/2}(X_i^0) = \left[\frac{1}{(n_i - 1)} \sum_{x_p \in X_i} \sum_{b=1}^{B} (\mu_{ib}^0 - \mu_{ib})^2 \right]^{1/2} \quad (5)$$

where μ_{ib}^0 is the mean for the b^{th} spectral band of region X_i at hierarchical level 0.

Spatial Features

A number of shape analysis measurements can be used to analyze the spatial properties of regions at the individual levels of detail in the segmentation hierarchy. The feature measurements considered in this study include 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*) and *extent*, defined as the proportion of the pixels in the bounding box (the smallest rectangle containing the region) that are also in the region. A major reason for selecting the four spatial-based metrics above was that they are widely available as built-in functions in several standard software products for remote sensing data and image analysis, including Research Systems, Inc.'s ENVI and Mathworks' Matlab packages.

Spectral Features

The spatial features in the previous section do not take into account the spectral information provided by MODIS and other multi- or hyperspectral data sets. To incorporate spectral signatures into automated selection of segmentation levels of detail, we used the Spectral Angle Mapper (SAM) and the Spectral Information Divergence (SID). While the SAM measure is popular and widely available in the remote sensing community, the use of SID for measuring spectral similarity in multspectral image data sets is new (Chang 2003). Consider two MODIS pixel vectors $x_i = (\chi_{i1}, \chi_{i2}, \ldots, \chi_{iB})^T$ and $x_j = (\chi_{j1}, \chi_{j2}, \ldots, \chi_{jB})^T$, where B is the number of spectral channels in the MODIS data. With the above definitions in mind, the SAM between x_i and x_j is given by:

$$\mathrm{SAM}(x_i, x_j) = \cos^{-1}\left(x_i \cdot x_j / \|x_i\| \|x_j\|\right) \qquad (6)$$

The SAM measurement 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, sensor observation angle, etc. In contrast, SID is based on the concept of divergence, and measures the discrepancy of probabilistic behaviors between two spectral signatures. It is based on a process that models x_i and x_j as random variables. Although this assumption does not necessarily hold true with remotely sensed images, the effects are negligible as explored by Chang (2003). If we assume that χ_{ib} and χ_{jb}, $b = 1, 2, \ldots, B$, are nonnegative values, which is a valid assumption for imagery data due to the nature of radiance and reflectance data, then two probability measures can be respectively defined for x_i and x_j as follows:

$$M[\chi_{ib}] = p_b = \chi_{ib} \Big/ \sum_{\beta=1}^{B} \chi_{i\beta}, \qquad M[\chi_{jb}] = q_b = \chi_{jb} \Big/ \sum_{\beta=1}^{B} \chi_{j\beta} \tag{7}$$

Using the above definitions, the self-information provided by x_j for band b is given by $I_b(x_j) = -\log q_b$. We can further define the relative entropy of x_j with respect to x_i, $D(x_i \| x_j)$ by:

$$D(x_i \| x_j) = \sum_{b=1}^{B} p_b D_b(x_i \| x_j) = \sum_{b=1}^{B} p_b (I_b(x_j) - I_b(x_i)) = \sum_{b=1}^{B} p_b \log(p_b / q_b) \tag{8}$$

By means of (8), a symmetric measure, referred to as SID, is defined as follows:

$$SID(x_i, x_j) = D(x_i \| x_j) + D(x_j \| x_i) \tag{9}$$

SID offers a new look at the spectral similarity between two spectral signatures by making use of relative entropy, and accounts for the spectral information provided by each spectral signature. Using the SAM and SID defined above; we can define a measure of spectral homogeneity for region X_i as:

$$S(X_i) = (1/n_i) \sum_{x_p \in X_i} \text{Dist}(x_p, u_i) \tag{10}$$

where n_i is the number of pixels in region X_i, x_p is a pixel vector (in this case, a pixel vector in region X_i), $u_i = (\mu_{i1}, \mu_{i2}, \ldots, \mu_{iB})^T$ is the mean vector of region X_i, Dist is either of SAM or SID. This measure provides indications of how similar are the spectral signatures of the pixel vectors labeled as part of the same region by RHSEG. Since the algorithm may associate pixels that are spatially disjoint but spectrally similar, the homogeneity measures should theoretically provide more reliable results than those produced by spatial-based metrics such as *solidity* or *extent*, both of which are addressed above.

Joint Spectral/Spatial Features

A combined spectral/spatial approach for feature extraction is based on classic mathematical morphology concepts (Soille 2003). Classic morphology utilizes two standard operations: erosion and dilation, which 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). The SE is a sliding window which is translated over all pixels in the input scene, thus acting as a probe for extracting or suppressing specific structures of the image objects by checking that each position of the SE fits within those objects. To extend the above two basic operations to multispectral images such as those produced by MODIS, we impose an ordering relation in the set of pixel vectors lying within an SE, designated by Z, by defining a cumulative distance between one particular pixel vector at spatial coordinates (x,y), denoted by $f(x,y)$, and all the pixel vectors in the spatial neighborhood given by Z (Z-neighborhood) as follows (Plaza et al. 2005):

$$D_Z[f(x,y)] = \sum_i \sum_j \text{Dist}[f(x,y), f(i,j)] \qquad (11)$$

where (i,j) refers to spatial coordinates in the Z-neighborhood and Dist refers either the SAM or SID distance. Based on the distance above, the extended erosion of f by Z selects the Z-neighborhood pixel vector that produces the minimum value for D_Z:

$$(f \ominus Z)(x,y) = \{f(x+i', y+j'), (i', j') = \arg\min_{(i,j)}\{D_Z[f(x+i, y+j)]\}\} \quad (12)$$

where the argmin operator selects the pixel vector that is most similar, spectrally, to all the other pixels in the Z-neighborhood. On other hand, the extended dilation of f by Z selects the Z-neighborhood pixel vector that produces the maximum value for D_Z:

$$(f \oplus Z)(x,y) = \{f(x-i', y-j'), (i', j') = \arg\max_{(i,j)}\{D_Z[f(x+i, y+j)]\}\} \quad (13)$$

where the argmax operator selects the pixel vector that is most spectrally distinct to all the other pixels in the Z-neighborhood. Based on the above operations, we define the Morphological Eccentricity Index (MEI) as a measure of spectral/spatial homogeneity at a given pixel as follows:

$$\text{MEI}(x,y) = \text{Dist}[(f \oplus Z)(x,y), (f \ominus Z)(x,y)] \quad (14)$$

where Dist may refer to either the SAM or SID distance. For both metrics we use the average MEI value for the pixels contained in region X_i as a measure of its spectral/spatial homogeneity the region.

RESULTS

Our long-term interest was in monitoring changes in land-use and land-cover. A study of this type will be facilitated by an ability to mask out confounding sections of the image data covered by land, clouds and/or snow and ice. We chose as an initial analysis task the generation of a land/water mask because having a good land/water mask made cloud mask generation much easier, since clouds can look spectrally different over water than they do over land. We chose the generation of snow/ice masks as our second task since the

analysis techniques involved are similar to, but somewhat simpler than, the general cloud mask generation problem. The MODIS standard product land/sea mask provided with the MOD03 geolocation data product was not applied because that this land/sea mask was too generalized for our purposes. We also chose not use the MODIS standard product snow/ice and cloud masks provided with the MOD35 product. With these masks we were concerned about the fine tuning required to set critical thresholds in the production of these products.

Land/Water Mask Generation

A quick look at the twelve MODIS data sets showed that none of them were completely cloud free. This complicated the task of generating a land/water mask. Nonetheless, a land/water mask was generated by the combining the analysis of all twelve MODIS data sets. Because of the transient nature of clouds, it was noted that the likelihood was high that any particular image location would be cloud-free for at least a small number of the images. Accordingly, RHSEG was used to generate segmentation hierarchies for each data set and each of these segmentation hierarchies was analyzed to find the super region associated with water. (A super region was a combination of individual spatially connected regions with similar spectral characteristics that may contain a number of spatially disjoint subregions.) Our goal is to find a measure that clearly designates the region at a particular segmentation hierarchy level that best characterizes this water super region for all data sets. We then will be able to combine the results from all twelve data sets to obtain the desired land/water mask.

Tables 3 (a, b) presents the minimum mean value region at selected hierarchical levels for each data set. For Table 3a, this information includes the number of pixels, the regon MMT_i, the region $BM\sigma_i$, the region $D^{1/2}_{BSMSE}(X_i)$ (vs. the original data values), and the region

$D^{1/2}_{BSMSE}(X_i^0)$ (vs. the finest hierarchical level (level 0)). Table 3b shows the results for the spatial, spectral and joint spatial/spectral measures described in the previous section.

As of version 1.15, RHSEG orders the regions from darkest (closest in vector distance to the zero vector) to brightest (furthest in vector distance from the zero vector). Since for MODIS bands 1-7, water usually is the darkest surface type, region 1 at the finest level of the segmentation hierarchy (level 0) consistently corresponded to water. This correspondence the darkest region to water also usually persisted through the coarsest segmentation hierarchy level. However, this was not the case for the March 20, 2004 and the June 11, 2004 data sets. Due to the influence of a large bank of clouds in the March 20, 2004 data set, the darkest region (region 20) at the coarsest level of the segmentation hierarchy (level 9) corresponded to a combination of water and land (Figure 1b). For the same reason, the darkest region (region 15) at the coarsest level of the segmentation hierarchy (level 13) corresponded to a combination of water and land in the June 11, 2004 data set (Figure 1c).

Table 3a the reports the region MMT_i, region $BM\sigma_i$, region $D^{1/2}_{BSMSE}(X_i)$, and region $D^{1/2}_{BSMSE}(X_i^0)$ all had relatively high values for the combined water and land regions as compared to regions that were purely water for both data sets. While these measures trended to higher values going from lower to higher hierarchical levels (that is, finer to coarser segmentations), the rise in values was much steeper where the region nature changed from water only to a combination of water and land. This was especially apparent in the March 20 data set for the region $D^{1/2}_{BSMSE}(X_i^0)$, with a value of 7.452 at hierarchical level 9. This measure dropped to well below 0.1 for level 8 of the segmentation hierarchy for the darkest region (region 1). From this point, the darkest region (region 1) remained constant from segmentation hierarchy level 8 down to level 3 and corresponded to water. This effect was

nearly as strong for the June 11 data set where the region $D^{1/2}_{BSMSE}(X_i^0)$, had a value of 5.077 for the combined water and land region. The measure dropped to about 0.3 for level 12 of the segmentation hierarchy for region 4 (the darkest region at hierarchical level 12). Region 4 remained constant from segmentation hierarchy level 12 down to level 4. At these hierarchical levels, this was the region that corresponded to water in this scene. The results indicated that low values of the region MMT_i, region $BM\sigma_i$, region $D^{1/2}_{BSMSE}(X_i)$ and region $D^{1/2}_{BSMSE}(X_i^0)$ serve as indicators of water, with the most sensitive measure being the region $D^{1/2}_{BSMSE}(X_i^0)$. All of these measures reflect the homogenous and uniform nature of the water regions.

Table 3b shows the values of the spatial features *solidity* and *extent* were relatively high for the March 20 data set at hierarchical level 9 (coarsest level), but dropped to very low values at the next finer hierarchical level, at which the region consisted solely of water. This was the only data set in which this phenomenon was observed. In contrast, the values of the spatial features *solidity* and *extent* were very high for the June data set at hierarchical level 13 (coarsest level), and did not change appreciably at the next finer hierarchical levels (consisting solely of water). Although it appeared that these measures might be useful in identifying some regions that changed appreciably from one hierarchical level to the next, other measures may be required to identify unidentified regions.

The values of the spectral homogeneity features SAM and SID (10) were very low for the March 20 and June 11 data at hierarchical levels 9 and 13, respectfully. However, they increased to very high values at the next finer hierarchical level, at which the region consisted solely of water. Since this trend was not observed for any of the other data sets, it appeared

that this trend is a moderate indicator of the change of the region nature from a combination of water and land to water only.

Similarly, the values of the spatial/spectral homogeneity features SAM and SID (14) were low for the March data set at hierarchical level 9, and for the June 11 data set at hierarchical level 13, but rose to very high values at the next finer hierarchical level, at which the region consisted solely of water. As with the spectral homogeneity feature, this trend was not observed for any of the other data sets with the spatial/spectral homogeneity features SAM and SID. It appeared that both versions of the SAM and SID features serve as moderate indicators of the change of the region nature from a combination of water and land to water only.

In summary, the spectral, spatial and spectral/spatial measures provided, at best, moderate indicator for determining at which hierarchical levels the segmentations consisted solely of water. However, some of these measures may prove to be more useful when looking for more structured objects, such as urban development and agricultural lands.

We found the region $D^{1/2}_{BSMSE}\left(X_i^0\right)$ to be a very strong indicator of the hierarchical level corresponding to the water region. Using this measure as our indicator, we created masks for each of the twelve data sets in which the value "1" corresponded to water and the value "0" corresponded to other (clouds, snow or land). These twelve masks were added to produce the composite mask shown in Figure 2. Inspection of the results showed that sum values 0 and 1 definitely corresponded to land and sum values 4 to 12 definitely corresponded to water. (Note that a cloud shadow could make a land pixel look like water in a particular scene.) Sum values 2 and 3 corresponded to pixels that may be water in some scenes and land in others, primarily along ocean and lake coastlines. Table 4 illustrates the minimum of the histogram corresponds to a natural threshold point distinguishing water from

land, where values equal to or higher than the histogram minimum correspond to water pixels (Figure 2).

Based on this analysis, pixels with sum values of 0 – 1 were set to land (green), pixels with sum values 2 – 3 were set to ocean coastlines and lake shorelines (red), and pixels with sum values of 4 – 12 were set to water (blue), creating the land/water mask (Figure 3). This mask can be visually compared to the MODIS standard product land/sea mask (Figure 4). Figure 4 illustrates deep and moderately deep ocean water and deep inland water (shades of blue), shallow ocean (white), ocean coastlines and lake shorelines (red), shallow inland water (turquoise), ephemeral (intermittent) water (gray) and land (green).

A quantitative comparison of the RHSEG-generated land/water mask to the MODIS standard product land/sea mask shows that the land and water areas were very similar in both cases (i.e., 600,339 vs. 595,019 km^2 for the land masks, and 403,665 vs. 377,244 km^2 for the water mask, respectively). Specifically, 98.5% of the image pixels labeled as land using RHSEG was also labeled as land in the MODIS standard product. For pixels labeled as water this figure dropped to 92.1%.

The above results indicate that the RHSEG generated land/water mask and the MODIS standard product land/sea mask may contain some relevant differences, especially in the water areas. In order to explore these differences more closely, a comparison of the RHSEG generated land/water mask and the MODIS standard product land/sea mask was made over a smaller area, as displayed in Figure 5. Here a 128 x 128 pixel section of data containing Lake Mead is shown at full resolution in Figure 5a. The corresponding RHSEG generated land/water mask is displayed in Figure 5b and the corresponding MODIS standard product land/sea mask is displayed in Figure 5c. The RHSEG generated land/water mask provided an estimated surface area of 609 km^2 for Lake Mead, while the MODIS standard

product land/sea mask indicated a much larger estimated surface area of 771 km^2. Further, only 59.5% of the pixels associated to the lake by the MODIS standard product were also associated to the lake by the RHSEG generated land/water mask. Since Lake Mead was near historical low levels during the study period, we believe that the MODIS standard product over-estimates Lake Mead's coverage. Being tailored to the particular study period, we believed that the land/water mask generated from RHSEG provides a more accurate estimate of Lake Mead's coverage for the study period.

From the above results, we conclude generally, that the MODIS standard product land/sea is a more conservative estimate of the land/water distinction with fuzzy categories such as ephemeral (intermittent) water and ocean coastlines and lake shorelines. As a result, this mask may be too generalized for detailed land cover, land use change monitoring. In contrast, the land/water mask generated with RHSEG provides a much crisper delineation of the land/water distinction that is tailored specifically to the data sets being analyzed.

Snow Mask Generation

Embedded within the MODIS Cloud Mask standard product is a snow/ice background flag. This flag is set based on a Normalized Snow Difference Index (NSDI) thresholded at a preset value (Hall, et al. 1995). NDSI uses the MODIS bands 4 and 6 to form a ratio where larger values correspond to snow or ice:

$$\text{NDSI} = \frac{R_4 - R_6}{R_4 + R_6} \tag{15}$$

where $R_{\#}$ is the reflectance value for band #. According to Hall, et al, (2002), a pixel in a non-densely forested region is mapped as snow/ice if the NDSI is ≥ 0.4, the MODIS band 2 reflectance was > 11%, and the MODIS band 4 reflectance is $\geq$ 10%. Application of this formula to the March 20, 2004 MODIS data set, and masking out water using the RHSEG

generate land/water mask from the previous section (with values 0 through 1 considered to be land), produced the snow/ice mask with 10,930 pixels were flagged as snow/ice (Figure 6a). Contrast this with the snow/ice background flag within the MODIS Cloud Mask standard product with only 5,316 pixels were flagged as snow/ice (Figure 6b). Clearly, the NDSI threshold for the MODIS standard product was adjusted upwards from 0.4 to avoid false positives (some clouds are flagged as snow/ice).

Through trial and error we found that a NDSI threshold of 0.68 produced a result with 5,304 pixels flagged as snow/ice. While very close in the number of pixels flagged, there were more than just 12 pixels flagged differently. Actually, there are 1,092 pixels labeled differently, with 540 false positives and 552 false negatives. These differences can be explained to a large degree by the fact that the MODIS standard product was calculated before georeferencing, while our NDSI 0.68 threshold result was calculated after georeferencing. Our NDSI 0.68 threshold result looks visually very similar to the MODIS standard product shown in Figure 6b.

When we applied RHSEG to the NDSI feature data from the March 20, 2004 data set, all negative NDSI values were set to zero and scaled by a factor of 10,000 before converting to unsigned short integer for input into RHSEG. Since RHSEG accepts only unsigned byte or unsigned short integer input data, and the NDSI feature is floating point, such a scaling scheme had to be adopted, since the masked NDSI feature ranged from -0.557 to 0.961. The processing mask was augmented to mask out all zero-valued pixels along with the previously masked out water pixels and pixels not satisfying the band 2 and band 4 thresholds. After masking, only 37,703 pixels out of 100,404 remained for RHSEG to process. As with the land/water mask generation task, RHSEG was run with the $d^{1/2}_{BSMSE}(X_i, X_j)$ dissimilarity criterion along with *spclust_wght* = 0.1, eight nearest neighbor connectivity, and the GMMT

convergence criterion. Again, our goal is to find a measure that clearly designates the region at a particular segmentation hierarchy level that best characterizes the snow/ice region.

The analysis results for the brightest two regions in the RHSEG generated segmentation hierarchy is given in Table 5 for our region and spatial features (the spectral and joint spectral/spatial features could not be calculated for the single band NSDI feature). The brightest region (region 38) containing 12,375 pixels at the coarsest hierarchical level (level 9) in the RHSEG output was clearly a mixed region due to the relatively high value (0.085) for the region $D^{1/2}_{BSMSE}(X_i^0)$ (Table 5). This region also has relatively high value for the region MMT_i and region $D^{1/2}_{BSMSE}(X_i)$. The values of the *solidity* and *extent* features are also highest for this region, but not as strongly as for the other features.

At hierarchical levels 5-8, the brightest region (region 59) contained 4871 pixels, compared to the 5316 pixels in the MODIS standard product. In this case the region $D^{1/2}_{BSMSE}(X_i^0)$ is fairly low at 0.10. The values of the region MMT_i and region $D^{1/2}_{BSMSE}(X_i)$ are also significantly lower than they were at hierarchical level 9. Also note that the mean value of the brightest region stays well above the 0.68 threshold we deduced above for the brightest region for hierarchical levels 0 through 8. These values led to conclude that selecting region 59 at hierarchical level 8 would be a conservative estimate of the snow/ice areas (Figure 6c). Comparing this region to the MODIS standard product reveals 1297 differently label pixels. Assuming the MODIS standard product was the "truth," the selected RHSEG region 59 contained 426 false positives and 871 false negatives.

Table 5 also shows that the next brightest region (region 55) at hierarchical levels 2 through 7 contains 2464 pixels. For this region the region $D^{1/2}_{BSMSE}(X_i^0)$ fell to a very low value of 0.001 at these hierarchical levels. The values of the region MMT_i and region

$D^{1/2}_{BSMSE}(X_i)$ are also relatively low, and the values of the *solidity* and *extent* measures are also lower, but not as strongly. The mean value of this region ranged from 0.63 through 0.66, just slightly below the previously deduced 0.68 threshold. These values led us to conclude that combining region 55 at hierarchical level 7 with the above selected region 59 at hierarchical level 8 would provide a somewhat more liberal 7335 pixel estimate of the snow/ice areas. This combined region is shown as the combination of the yellow and white areas in Figure 6c. Comparing this combined region to the MODIS standard product reveals 2223 differently label pixels. Assuming the MODIS standard product was the "truth," the selected combination of RHSEG regions 55 and 59 contained 2121 false positives and 102 false negatives.

We tried other values of *spclust_wght* in running RHSEG before settling on a value of 0.1. A value of 0.01 produced results very similar to that produced with the value 0.1. However, with a value of 1.0, RHSEG produced a brightest region at the coarsest hierarchical level consisting of 6572 pixels with a spatial distribution similar to the *spclust_wght* = 0.1 results with the 4871 pixel region 59 at hierarchical level 8. However, the next brightest region at the next coarsest hierarchical level included cloud pixels from a bright cloud in the southeast part of the data set. Such confusion with clouds did not occur in the *spclust_wght* = 0.1 results with region 55 at hierarchical level 7.

To examine the behavior of RHSEG in more detail, we present in Figure 7 a portion of the segmentation hierarchy for the *spclust_wght* = 0.1 results. At hierarchical level 0 there were 64 regions ordered from darkest to brightest and at level 1 several of these regions were merged. In the portion of the segmentation hierarchy shown in Figure 7, region 64 merged into region 62, regions 60, 57, 56, 54, 53, 52, 50 and 49 merged into region 55, and regions 51, 48, 47 and a number of other regions merged into region 43. The region merges were not

necessarily into the region with the closest mean value due to the influence of *spclust_wght*. For example, region 64's mean value was closer to that of region 63, but it instead merged into region 62 because it was spatially adjacent to region 62. Likewise, region 60 merged into region 55 instead of region 61 (the closest region in mean value) because it was spatially adjacent to region 55. Other such influences of *spclust_wght* can be found in the structure of the segmentation hierarchy.

This influence of *spclust_wght* is the reason why the *spclust_wght* = 0.1 results with the selected combination of region 55 and 59 found 7335 snow/ice pixels without confusion with clouds; while the going beyond selecting the brightest region with 6572 pixels at the coarsest hierarchical level led to confusion with clouds. Since with *spclust_wght* = 0.1, a 10 times priority is given to merges between spatially adjacent regions, slightly less bright snow/ice areas adjacent to brighter snow/ice areas were merged into the adjacent brighter snow/ice areas rather than into nonadjacent cloud areas that may have had more similar region mean values.

Figure 8 provides a more detailed comparison of the snow/ice mask detected by RHSEG and the MODIS standard product snow/ice background flag, along with a view of the MODIS data itself. For the most part, the selection of RHSEG region 59 at hierarchical level was very similar to the MODIS standard product (Figs. 6b and 8c). The main exception to this was the snow field in the lower portion of Figure 8. To produce a snow/ice mask similar to the MODIS standard product, we needed to add RHSEG region 55 at hierarchical level.

From this analysis we cannot determine which snow/ice mask is more accurate. To determine this we would have to work with the MODIS production team to perform studies in which we have actual ground truth. We intend to propose such an effort to the MODIS

production team using this paper to introduce the team to our approach. Until such a definitive study is performed we can only speculate as to the significance of our results. We hypothesize that selecting the brightest RHSEG region at the coarsest hierarchical level consistent with region coherence (most strongly indicated by low values of the region $D^{1/2}_{BSMSE}(X_i^0)$) will provide a consistent conservative estimate of the snow/ice area. A more liberal estimate could be obtained by adding a next brighter region.

CONCLUSIONS

Here we explored an approach for the automatic generation of land/water and snow/ice masks from MODIS data utilizing segmentation hierarchies produced by RHSEG, the recursive approximation of the hierarchical segmentation algorithm (HSEG). These segmentation hierarchies organize the image data in a manner that makes the image's information content more accessible for analysis by enabling region-based analysis. Further, the segmentation hierarchies provided additional analysis clues through the behavior of the image region characteristics over several levels of segmentation detail.

A main contribution of this paper was the introduction of several region features for characterizing the nature of regions in the segmentation hierarchy. A simple region dissimilarity measure was seen to be effective in identifying water regions and snow/ice regions. This measure was the region square root of the Band Sum Mean-Squared Error vs. the region mean value at the finest hierarchical level, $D^{1/2}_{BSMSE}(X_i^0)$. Other region feature measures were not as consistent or distinctive in such identification. However, we also observed that a detailed post-analysis of segmentation hierarchies through the utilization of spatial and/or spectral metrics provided additional clues that be used for the purpose of automatically selecting regions out of the segmentation hierarchy. While these measures

were not as useful for identifying water and snow/ice regions, they may prove to be more useful when looking for more structured objects, such as urban development and agricultural lands. This constitutes a novel approach to extracting information from multi-temporal remotely sensed image data.

Experimental results presented in this paper explored measures of region characteristics for use in the automatic selection of appropriate regions from segmentation hierarchies corresponding to water, and to snow and ice. We showed how this approach can be used to generate a land/water mask from a set of twelve multi-temporal data sets collected by MODIS. We then used similar techniques to generate a snow/ice mask for a selected MODIS data. The RHSEG generated land/water and snow/ice masks were compared with similar standard products provided with the MODIS data. Interestingly, we found that our RHSEG-generated masks compared very favorably with those available from MODIS. Based on our previous knowledge about the seasonal properties of the Lake Mead area, we believe that the RHSEG-based approach can provide more accurate land/water estimates for a particular study period. We also believe that the RHSEG-based approach can provide more complete snow/ice masks without confusion with clouds. However, further experimentation with additional data sets with detailed ground truth is required to fully validate these conclusions.

Our motivation for accurate generation of land/water and snow/ice masks is our long term goal to use these masks along with similarly generated cloud masks to facilitate the monitoring of land-cover and land-use change using multi-spectral remotely sensed imagery. As an additional step toward our long term goal, we explored various spectral and region features that could be used in the analysis of segmentation hierarchies for land-cover change,

and noted the behavior of these measures relative to the land/water and snow/ice mask generation tasks.

REFERENCES

Beaulieu, J.-M. and M. Goldberg, 1989, "Hierarchy in Picture Segmentation: A Stepwise Optimal Approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(2):150-163.

Beaulieu, J.-M., 2004, "Utilization of contour criteria in micro-segmentation of SAR images," *International Journal of Remote Sensing*, 25(17):3497-3512.

Chang, C.-I, 2003, *Hyperspectral Imaging: Techniques for Spectral Detection and Classification*, Kluwer Academic/Plenum Publishers: New York.

Hall, D. K., G. A. Riggs, and V. V. Salomonson, 1995, "Development of methods for mapping global snow cover using Moderate Resolution Imaging Spectroradiometer data," *Remote Sensing of the Environment*, 54:127-140.

Hall, D. K., G. A. Riggs, V. V. Salomonson, N. E. DiGirolamo, and K. J. Bayr, 2002, "MODIS snow-cover products," *Remote Sensing of Environment*, 83:181-194.

Plaza, A., P. Martínez, J. Plaza, and R. Pérez, 2005, "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.

Soille, P., 2003, *Morphological Image Analysis: Principles and Applications*, 2nd Ed., Springer-Verlag: New York.

Tilton, J. C. and S. C. Cox, 1983, "Segmentation of Remotely Sensed Data using Parallel Region Growing," *1983 International Geoscience and Remote Sensing Symposium (IGARSS'83) Digest*, San Francisco, CA, Vol. 1, Section WP-4, paper 9.

Tilton, J. C., 1998, "Image Segmentation by Region Growing and Spectral Clustering with a Natural Convergence Criterion," *Proceedings of the 1998 International Geoscience and Remote Sensing Symposium*, Seattle, WA, pp. 1766-1768.

TABLE AND FIGURE CAPTIONS

Table 1. Summary of Acronyms and Notation.

Table 2. MODIS Data Sets Studied. All data granules are from the Terra, AM1 Platform, which has a 705 Km Sun-Synchronous Near-Polar Circular Orbit with a 10:30am Equatorial Crossing in its Descending Node.

Table 3a. Analysis of the Darkest Region from the Segmentation Hierarchies of each of the Twelve MODIS Data Sets using Region Features.

Table 3b. Analysis of the Darkest Region from the Segmentation Hierarchies of each of the Twelve MODIS Data Sets using Spatial, Spectral and Spatial/Spectral Features.

Table 4. Histogram of Figure 2.

Table 5. Analysis of the RHSEG Segmentation Hierarchy of the Brightest Region and Next Brightest Region for the Selection of the Hierarchical Level Corresponding to Snow and Ice.

Figure 1. Four of the twelve MODIS data sets utilized in the study. Reflective bands 7, 2 and 1 are displayed as red, green and blue, respectively. (a) The October 28, 2003 image shows smoke wafting over the ocean from large chaparral fires along the southern California coast. (b) The March 20, 2004 image shows a large bank of clouds over most of the ocean and some snow in the mountains in the north central and north east areas. (c) The June 11, 2004 image shows a large bank of clouds over the south west ocean area. (d) The February 28, 2005 image shows a minor amount of haze over the ocean and snow in the mountains in the north central and north east areas. This data set is also cut-off over a part of the ocean in the southwest corner of the image.

Figure 2. Sum of the twelve individual RHSEG generated water masks from the multi-temporal data set. Even though some areas of the ocean look fairly dark, the sum value at those locations is still high enough to indicate water (see Figure 3).

Figure 3. Land/water mask created from Figure 2 by setting values 0-1 to land (green), values 2-3 to ocean coastlines and lake shorelines (red), and values 4-12 as water (blue).

Figure 4. MODIS standard land/sea mask product. Various shades of blue are deep and moderately deep ocean water and deep inland water, white is shallow ocean, red is ocean coastlines and lake shorelines, turquoise is shallow inland water, gray is ephemeral (intermittent) water and green is land.

Figure 5. Comparison of a portion of the land/water mask generated from RHSEG with the land/sea mask from the MODIS standard product MOD3 for Lake Mead, NV. Since Lake Mead was near historical low levels during the study period, the MODIS standard product over-estimates Lake Mead's coverage. Being tailored to the particular study period, the land/water mask generated from RHSEG is provides a more accurate estimate of Lake Mead's coverage for the study period (a) October 28, 2003 MODIS image over Lake Mead, (b) Land/water mask generated from RHSEG, and (c) Land/sea mask from MODIS standard product MOD3.

Figure 6. Snow/ice areas designated in yellow as detected by (a) thresholding NDSI at 0.4, (b) from the MODIS standard product, and (c) detected by RHSEG by selecting the brightest region at the second coarsest hierarchical level. The white areas designate additions to the snow/ice mask found by also selecting the second brightest region at the third coarsest hierarchical level.

Figure 7. A portion of the segmentation hierarchy for the RHSEG *spclust_wght* = 0.1 results for the snow/mask generation task. At each node in the segmentation hierarchy is labeled with either the region label alone (unchanged), or with the region label, number of pixels and region mean (label:pixels:mean). The arrows indicate the merging hierarchy. It is significant to note that regions do not necessarily merge with the most similar region in mean value. With *spclust_wght* = 0.1, spatial connectivity influences merge behavior.

Figure 8. A detailed comparison of the snow/ice mask detected by (a) MODIS composite image (bands 6, 4, 2), (b) RHSEG, and (c) the MODIS standard snow/ice product. The yellow area, obtained by selecting the brightest region at the second coarsest hierarchical level, was very similar to the snow/ice mask provided by the MODIS standard product. An exception to this was the snow field in the lower right. Here we need to add RHSEG region 55 at hierarchical level 7 (the white area) to obtain a snow/ice mask similar to the MODIS standard product.

APPENDIX A

The square root of Band Sum Mean Squared Error dissimilarity function is based on minimizing the increase of mean squared error (MSE) between the region mean image and the original image data. The sample estimate of the MSE for the segmentation of band b of the image X into R disjoint subsets (regions) $X_1, X_2, \ldots, X_R$ is given by:

$$\mathrm{MSE}_b(X) = \frac{1}{N-1}\sum_{i=1}^{R}\mathrm{MSE}_b(X_i), \tag{A-1a}$$

where N is the total number of pixels in the image data set and

$$\mathrm{MSE}_b(X_i) = \sum_{x_p \in X_i}(\chi_{pb} - \mu_{ib})^2 \tag{A-1b}$$

is the MSE contribution for band b from data subset or region X_i. Here, x_p is a pixel vector (in this case, a pixel vector in region X_i), χ_{pb} is the image data value for the b^{th} spectral band of the pixel vector, x_p, and μ_{ib} is the mean for spectral band b of data region X_i. A dissimilarity function based on a measure of the increase in MSE due to the merge of regions X_i and X_j is given by:

$$\mathrm{d}_{\mathrm{BSMSE}}(X_i, X_j) = \sum_{b=1}^{B}\Delta\mathrm{MSE}_b(X_i, X_j), \tag{A-2a}$$

where

$$\Delta\mathrm{MSE}_b(X_i,X_j) = \mathrm{MSE}_b(X_i \cup X_j) - \mathrm{MSE}_b(X_i) - \mathrm{MSE}_b(X_j)\ . \tag{A-2b}$$

BSMSE refers to Band Sum MSE.

Using (A-1b) and exchanging the order of summation, (A-2b) can be manipulated to produce an efficient dissimilarity function based on aggregated region features:

$\Delta\mathrm{MSE}_b(X_i,X_j) =$

$$\left(\sum_{x_p \in X_{ij}}[\chi_{pb} - \mu_{ijb}]^2 - \sum_{x_p \in X_i}[\chi_{pb} - \mu_{ib}]^2 - \sum_{x_p \in X_j}[\chi_{pb} - \mu_{jb}]^2\right) =$$

$$\left(\sum_{x_p \in X_i}\left[(x_{pb}-\mu_{ijb})^2-(x_{pb}-\mu_{ib})^2\right]+\sum_{x_p \in X_j}\left[(x_{pb}-\mu_{ijb})^2-(x_{pb}-\mu_{jb})^2\right]\right)=$$

(algebraic manipulations)

$$n_i(\mu_{ib}-\mu_{ijb})^2+n_j(\mu_{jb}-\mu_{ijb})^2. \qquad \text{(A-3a)}$$

where μ_{ijb} is the mean value for the b^{th} spectral band of the mean vector, u_{ij}, of region represented by $X_{ij} = X_i \cup X_j$, and n_i (n_j) is the number of pixels in region X_i (X_j).

Since

$$\mu_{ijb} = \frac{n_i\mu_{ib} + n_j\mu_{jb}}{n_i + n_j}, \qquad \text{(A-3b)}$$

an alternate form for (A-3a) is:

$\Delta MSE_b(X_i,X_j) =$

$$n_i(\mu_{ib}-\mu_{ijb})^2+n_j(\mu_{jb}-\mu_{ijb})^2 =$$

$$n_i\mu_{ib}^2+n_j\mu_{jb}^2-2(n_i\mu_{ib}+n_j\mu_{jb})\mu_{ijb}+(n_i+n_j)\mu_{ijb}^2$$

$$\frac{1}{(n_i+n_j)}\left[(n_i+n_j)(n_i\mu_{ib}^2+n_j\mu_{jb}^2)-2(n_i\mu_{ib}+n_j\mu_{jb})^2+(n_i\mu_{ib}+n_j\mu_{jb})^2\right]=$$

$$\frac{1}{(n_i+n_j)}\left(n_i^2\mu_{ib}^2+n_in_j\mu_{jb}^2+n_in_j\mu_{ib}^2+n_j^2\mu_{jb}^2-n_i^2\mu_{ib}^2-2n_in_j\mu_{ib}\mu_{jb}-n_j^2\mu_{jb}^2\right)=$$

$$\frac{n_in_j}{(n_i+n_j)}(\mu_{ib}-\mu_{jb})^2. \qquad \text{(A-3c)}$$

Combining (A-2a) and (A-3c),

$$d_{BSMSE}(X_i,X_j)=\frac{n_in_j}{(n_i+n_j)}\sum_{b=1}^{B}(\mu_{ib}-\mu_{jb})^2. \qquad \text{(A-4)}$$

The dimensionality of the $d_{BSMSE}(X_i,X_j)$ dissimilarity criterion is equal to the square of the dimensionality of the image pixel values, while the dimensionality of the vector norm based dissimilarity criteria is equal to the dimensionality of the image pixel values. To keep

the dissimilarity criteria dimensionality consistent, HSEG and RHSEG use the square root of the d_{BSMSE} dissimilarity criterion. Thus, the square root of Band Sum Mean Squared Error criterion, $d^{1/2}_{BSMSE}(X_i, X_j)$, is:

$$d^{1/2}_{BSMSE}(X_i, X_j) = \left[\frac{n_i n_j}{(n_i + n_j)} \sum_{b=1}^{B} (\mu_{ib} - \mu_{jb})^2 \right]^{1/2}. \qquad \text{(A-5)}$$

TABLES

TABLE 1. SUMMARY OF ACRONYMS AND NOTATION.

BSQ	Band Sequential (data format)
GMMT	Global Maximum Merge Threshold encountered for all regions through the HSEG region growing process
HDF	Hierarchical Data Format
HSEG	Hierarchical Segmentation
HSEGViewer	Program for visualizing and manipulating HSEG segmentation hierarchy results
HSWO	Hierarchical Step-Wise Optimization
MEI	Morphology Eccentricity Index (14)
MMT_i	Maximum Merge Threshold encountered for region i through the HSEG region growing process
MOD021KM	MODIS/Terra Level 1B calibrated and geolocated radiances at-aperture at 1km resolution
MOD03	MODIS/Terra Geolocation including geodetic latitude and longitude and other information
MOD35	Level 2 MODIS/Terra cloud mask product
MODIS	Moderate Resolution Imaging Spectroradiometer
MSE	Mean Squared Error
NSDI	Normalized Snow Difference Index (15)
RHSEG	Recursive approximation of HSEG
SAM	Spectral Angle Mapper
SAR	Synthetic Aperture Radar
SE	Structuring Element
SID	Spectral Information Divergence
$BM\sigma_i$	Band Maximum Standard Deviation of region i (3)
$d_{BSMSE}(X_i, X_j)$	Band Sum Mean Squared Error dissimilarity criterion between regions i and j (A-2a) (A-4)
$d^{1/2}_{BSMSE}(X_i, X_j)$	Square root of Band Sum Mean Squared Error dissimilarity criterion between regions i and j (1) (A-5)
$D^{1/2}_{BSMSE}(X_i)$	Square root of the Band Sum Mean Squared Error dissimilarity criterion for region i (versus original image data) (4)
$D^{1/2}_{BSMSE}(X^0_i)$	Square root of the Band Sum Mean Squared Error dissimilarity criterion for region i versus the region mean at segmentation hierarchy level 0 (5)
$MSE_b(X)$	Mean Squared Error for the segmentation of band b of the image X (A-1a)
$MSE_b(X_i)$	Mean Squared Error for the segmentation of band b of region i of image X (A-1b)
$\Delta MSE_b(X_i,X_j)$	Increase in MSE due to the merge of regions i and j (A-2b) (A-3a) (A-3c)
B	Number of spectral bands
R	Number of regions (in an image)
X_i	Image data subset or segment for region i
n_i	Number of pixels in region i
x_p	Image pixel vector at location p
x_{pb}	Image data value for the b^{th} spectral band of pixel vector x_p
u_i	Mean vector for region i
u_{ib}	Mean value for the b^{th} spectral band of region i
μ^0_{ib}	Mean for the b^{th} spectral band of region i at segmentation hierarchy level 0
u_{ijb}	Mean value for the b^{th} spectral band of the union of regions i and j (A-3b)
σ_{ib}	Standard deviation for spectral band b of region i (2)

TABLE 2. MODIS DATA SETS STUDIED. ALL DATA GRANULES ARE FROM THE TERRA, AM1 PLATFORM, WHICH HAS A 705 KM SUN-SYNCHRONOUS NEAR-POLAR CIRCULAR ORBIT WITH A 10:30AM EQUATORIAL CROSSING IN ITS DESCENDING NODE.

MOD03 & MOD021KM Data Granules		
Granule name	***Date***	***Notes***
A2003031.1815	31 JAN 2003	Winter
A2003109.1825	19 APR 2003	Spring
A2003221.1825	09 AUG 2003	Summer
A2003294.1820	21 OCT 2003	Pre-fire
A2003301.1825	28 OCT 2003	Fire
A2003322.1845	18 NOV 2003	Post-fire
A2004032.1825	01 FEB 2004	Winter
A2004080.1825	20 MAR 2004	Spring
A2004163.1855	11 JUN 2004	Summer
A2004268.1850	24 SEP 2004	Summer
A2004334.1835	29 NOV 2004	Winter
A2005059.1820	28 FEB 2005	Spring

TABLE 3A. ANALYSIS OF THE DARKEST REGION FROM THE SEGMENTATION HIERARCHIES OF EACH OF THE TWELVE MODIS DATA SETS USING REGION FEATURES.

Data Set	Region Label	Hier-archical Levels	Number of Pixels	MMT_i	Region $BM\sigma_i$	Region $D^{1/2}_{BSMSE}(X_i)$	Region $D^{1/2}_{BSMSE}(X_i^0)$
31 JAN	3	6-9	372536	245.9	0.280	0.601	0.279
2003	3	4-5	368493	180.9	0.247	0.441	0.250
	3	2-3	269463	84.7	0.182	0.256	0.120
	1	0-1	48808	15.2	0.085	0.188	0.000
19 APR	1	6-9	278756	221.3	0.257	0.556	0.156
2003	1	1-5	171688	40.4	0.092	0.181	0.006
	1	0	105167	16.4	0.056	0.127	0.000
09 AUG	1	9-11	352823	290.7	0.333	0.656	0.141
2003	1	1-8	232608	30.9	0.049	0.104	0.001
	1	0	204260	16.8	0.047	0.080	0.000
21 OCT	1	6-12	238404	117.4	0.140	0.318	0.012
2003	1	2-5	226660	40.3	0.102	0.158	0.005
	1	0-1	134909	21.0	0.068	0.151	0.000
28 OCT	2	7-8	398465	221.2	0.263	0.467	0.359
2003	2	3-6	202587	65.6	0.143	0.230	0.074
	1	0-2	45386	13.8	0.089	0.197	0.000
18 NOV	1	5-9	331974	138.3	0.150	0.303	0.012
2003	1	1-4	312726	31.8	0.068	0.143	0.003
	1	0	170081	4.8	0.035	0.078	0.000
01 FEB	2	3-10	286303	113.3	0.165	0.317	0.048
2004	2	1-2	250221	55.9	0.131	0.205	0.024
	1	0	84518	19.4	0.073	0.168	0.000
20 MAR	20	9	688461	821.9	0.698	1.443	7.452
2004	1	3-8	89699	78.3	0.152	0.377	0.035
	1	0-2	59152	13.7	0.075	0.182	0.000
11 JUN	15	13	501835	703.1	0.642	1.269	5.077
2004	4	5-12	124884	125.9	0.225	0.473	0.302
	2	1-4	62398	42.3	0.138	0.319	0.038
	1	0	26790	0.0	0.028	0.049	0.000
24 SEP	3	2-8	355922	59.4	0.082	0.182	0.044
2004	1	1	44561	26.0	0.126	0.194	0.002
	1	0	40352	11.1	0.064	0.128	0.000
29 NOV	1	3-13	389655	82.6	0.117	0.235	0.007
2004	1	2	383441	61.4	0.097	0.184	0.005
	1	1	375235	39.0	0.066	0.142	0.003
	1	0	215663	14.2	0.037	0.070	0.000
28 FEB	3	7-11	286362	148.3	0.187	0.394	0.159
2005	3	2-6	210300	56.5	0.143	0.209	0.070
	1	0-1	37718	0.0	0.047	0.108	0.000

TABLE 3B. ANALYSIS OF THE DARKEST REGION FROM THE SEGMENTATION HIERARCHIES OF EACH OF THE TWELVE MODIS DATA SETS USING SPATIAL, SPECTRAL AND SPATIAL/SPECTRAL FEATURES.

Data Set	Region Label	Hier-archical Levels	*solidity*	*extent*	Spectral Homogeneity		Spatial/Spectral Homogeneity	
					SAM	SID	SAM	SID
31 JAN	3	6-9	0.816	0.653	0.986	0.958	0.891	0.873
2003	3	4-5	0.785	0.594	0.971	0.959	0.878	0.865
	3	2-3	0.678	0.499	0.925	0.912	0.823	0.819
	1	0-1	0.723	0.589	0.962	0.945	0.837	0.824
19 APR	1	6-9	0.784	0.729	0.981	0.972	0.903	0.891
2003	1	1-5	0.798	0.740	0.978	0.969	0.925	0.909
	1	0	0.756	0.717	0.965	0.958	0.912	0.901
09 AUG	1	9-11	0.892	0.827	0.974	0.962	0.938	0.926
2003	1	1-8	0.761	0.742	0.951	0.942	0.890	0.881
	1	0	0.776	0.754	0.948	0.931	0.887	0.874
21 OCT	1	6-12	0.725	0.592	0.928	0.918	0.834	0.899
2003	1	2-5	0.718	0.580	0.917	0.906	0.823	0.811
	1	0-1	0.696	0.548	0.925	0.916	0.794	0.785
28 OCT	2	7-8	0.944	0.885	0.934	0.919	0.933	0.928
2003	2	3-6	0.803	0.712	0.912	0.906	0.883	0.865
	1	0-2	0.724	0.643	0.909	0.901	0.824	0.817
18 NOV	1	5-9	0.889	0.828	0.934	0.925	0.880	0.779
2003	1	1-4	0.862	0.813	0.928	0.917	0.871	0.865
	1	0	0.789	0.698	0.932	0.920	0.801	0.794
01 FEB	2	3-10	0.777	0.701	0.953	0.941	0.876	0.867
2004	2	1-2	0.735	0.640	0.941	0.928	0.823	0.808
	1	0	0.724	0.623	0.932	0.921	0.812	0.807
20 MAR	20	9	0.622	0.567	0.589	0.561	0.589	0.568
2004	1	3-8	0.385	0.343	0.875	0.853	0.776	0.765
	1	0-2	0.394	0.358	0.882	0.874	0.791	0.782
11 JUN	15	13	0.612	0.684	0.574	0.562	0.625	0.612
2004	4	5-12	0.645	0.712	0.941	0.933	0.923	0.915
	2	1-4	0.699	0.723	0.956	0.940	0.919	0.908
	1	0	0.678	0.702	0.974	0.972	0.912	0.903
24 SEP	3	2-8	0.868	0.809	0.951	0.946	0.932	0.920
2004	1	1	0.733	0.649	0.971	0.962	0.845	0.826
	1	0	0.721	0.636	0.967	0.954	0.823	0.810
29 NOV	1	3-13	0.950	0.829	0.972	0.969	0.942	0.930
2004	1	2	0.938	0.813	0.973	0.968	0.931	0.924
	1	1	0.934	0.802	0.971	0.962	0.925	0.917
	1	0	0.824	0.695	0.964	0.956	0.909	0.900
28 FEB	3	7-11	0.789	0.671	0.952	0.940	0.931	0.920
2005	3	2-6	0.615	0.566	0.948	0.942	0.884	0.872
	1	0-1	0.489	0.508	0.931	0.929	0.849	0.830

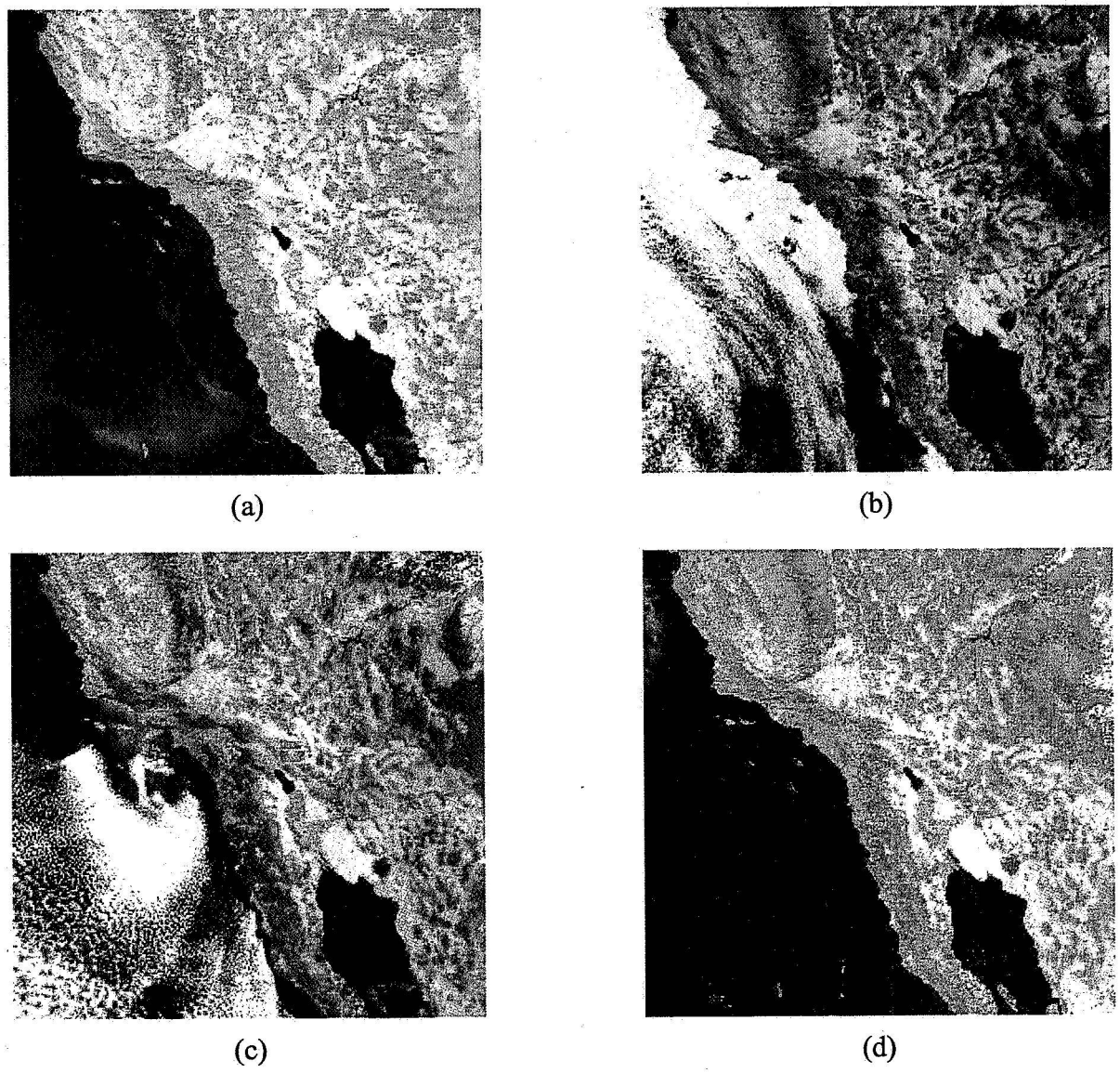
TABLE 4. HISTOGRAM OF FIGURE 2.

Value	Occurrences
0	600339
1	2097
2	940
3	849
4	844
5	3499
6	20311
7	62839
8	107174
9	83883
10	59408
11	31616
12	30205

TABLE 5. ANALYSIS OF THE RHSEG SEGMENTATION HIERARCHY OF THE BRIGHTEST REGION AND NEXT BRIGHTEST REGION FOR THE SELECTION OF THE HIERARCHICAL LEVEL CORRESPONDING TO SNOW AND ICE.

Region Label	Hier-archical Levels	Number of Pixels	Region Mean	MMT_i	Region $D^{1/2}_{BSMSE}(X_i)$	Region $D^{1/2}_{BSMSE}(X_i^0)$	*solidity*	*extent*
Brightest:								
38	9	12375	0.63	17.09	0.188	0.085	0.816	0.692
59	5-8	4871	0.82	3.05	0.077	0.010	0.725	0.631
59	4	3624	0.79	2.23	0.071	0.016	0.654	0.503
62	1-3	555	0.88	0.60	0.041	0.001	0.779	0.559
64	0	170	0.92	0.00	0.024	0.000	0.762	0.536
Next Brightest:								
38	9	12375	0.63	17.09	0.188	0.085	0.798	0.658
38	8	7504	0.50	7.46	0.125	0.025	0.734	0.526
55	2-7	2464	0.63	1.32	0.080	0.001	0.699	0.492
55	1	2193	0.64	0.99	0.074	0.001	0.694	0.487
55	0	616	0.66	0.40	0.065	0.000	0.718	0.495

FIGURES



(a) (b)

(c) (d)

Figure 1. Four of the twelve MODIS data sets utilized in the study. Reflective bands 7, 2 and 1 are displayed as red, green and blue, respectively. (a) The October 28, 2003 image shows smoke wafting over the ocean from large chaparral fires along the southern California coast. (b) The March 20, 2004 image shows a large bank of clouds over most of the ocean and some snow in the mountains in the north central and north east areas. (c) The June 11, 2004 image shows a large bank of clouds over the south west ocean area. (d) The February 28, 2005 image shows a minor amount of haze over the ocean and snow in the mountains in the north central and north east areas. This data set is also cut-off over a part of the ocean in the southwest corner of the image.

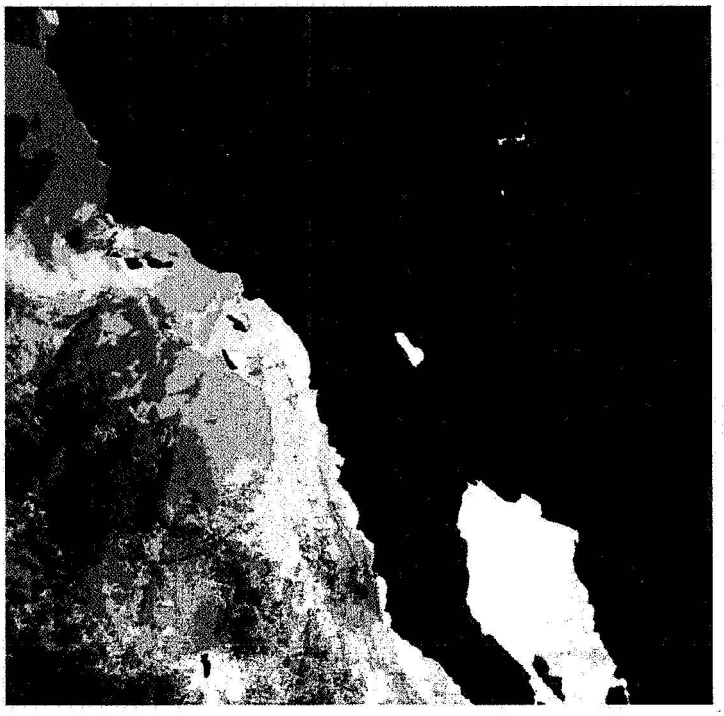


Figure 2. Sum of the twelve individual RHSEG generated water masks from the multi-temporal data set. Even though some areas of the ocean look fairly dark, the sum value at those locations is still high enough to indicate water (see Figure 3).

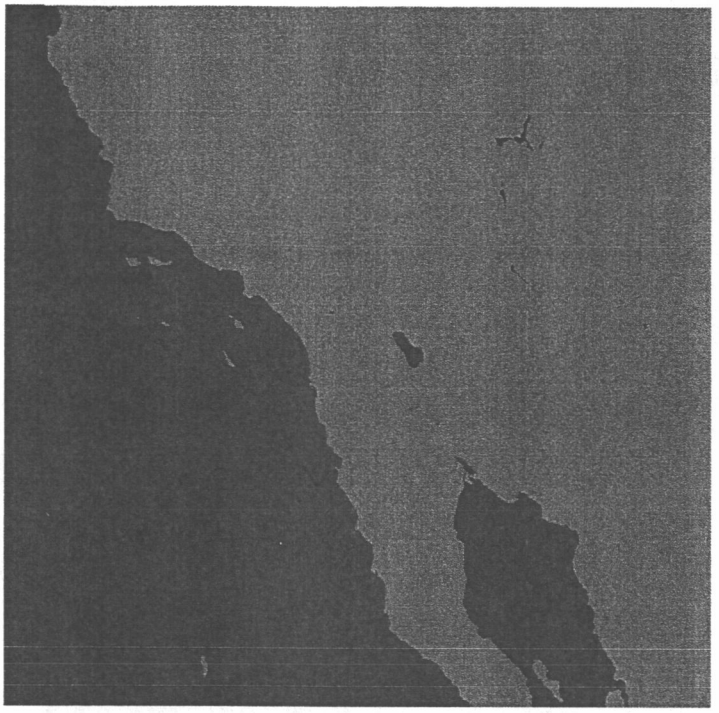


Figure 3. Land/water mask created from Figure 2 by setting values 0-1 to land (green), values 2-3 to ocean coastlines and lake shorelines (red), and values 4-12 as water (blue).

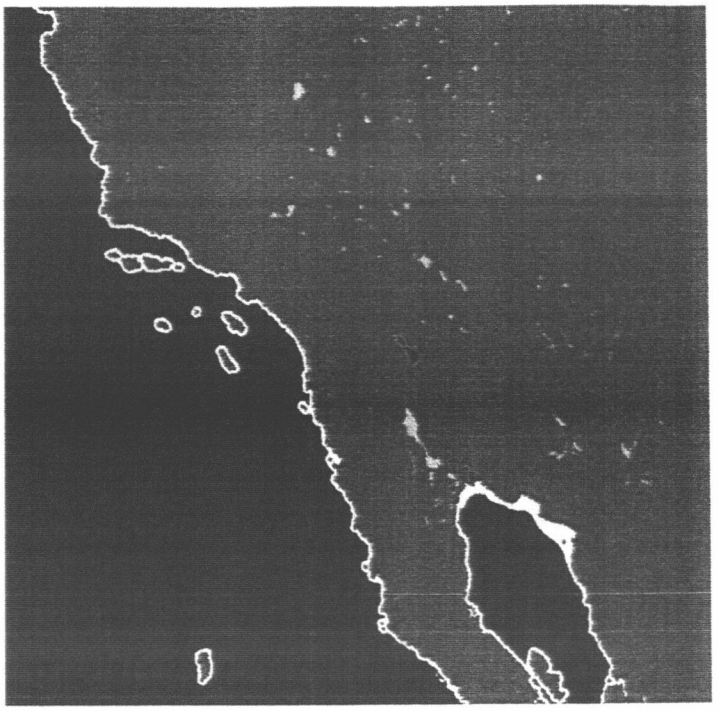
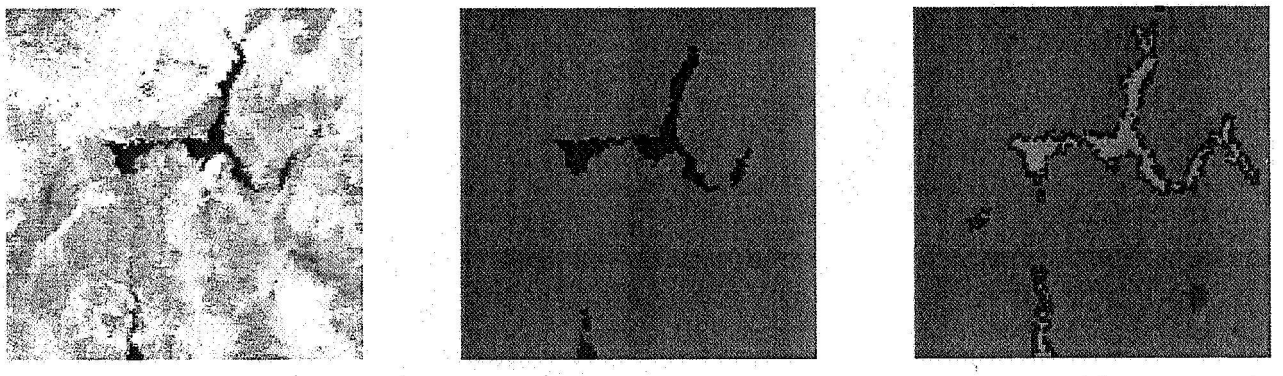


Figure 4. MODIS standard land/sea mask product. Various shades of blue are deep and moderately deep ocean water and deep inland water, white is shallow ocean, red is ocean coastlines and lake shorelines, turquoise is shallow inland water, gray is ephemeral (intermittent) water and green is land.

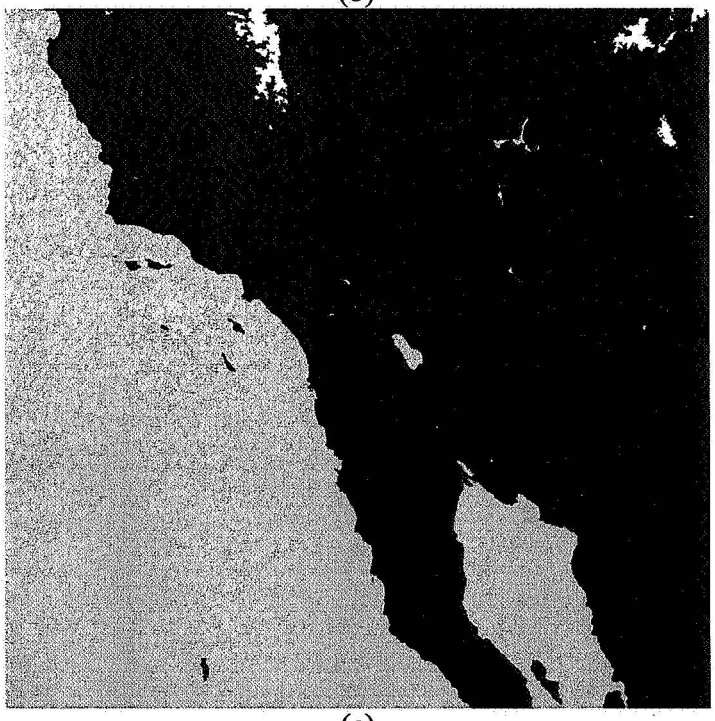


(a) (b) (c)

Figure 5. Comparison of a portion of the land/water mask generated from RHSEG with the land/sea mask from the MODIS standard product MOD3 for Lake Mead, NV. Since Lake Mead was near historical low levels during the study period, the MODIS standard product over-estimates Lake Mead's coverage. Being tailored to the particular study period, the land/water mask generated from RHSEG is provides a more accurate estimate of Lake Mead's coverage for the study period (a) October 28, 2003 MODIS image over Lake Mead, (b) Land/water mask generated from RHSEG, and (c) Land/sea mask from MODIS standard product MOD3.

(a)

(b)



(c)

Figure 6. Snow/ice areas designated in yellow as detected by (a) thresholding NDSI at 0.4, (b) from the MODIS standard product, and (c) detected by RHSEG by selecting the brightest region at the second coarsest hierarchical level. The white areas designate additions to the snow/ice mask found by also selecting the second brightest region at the third coarsest hierarchical level.

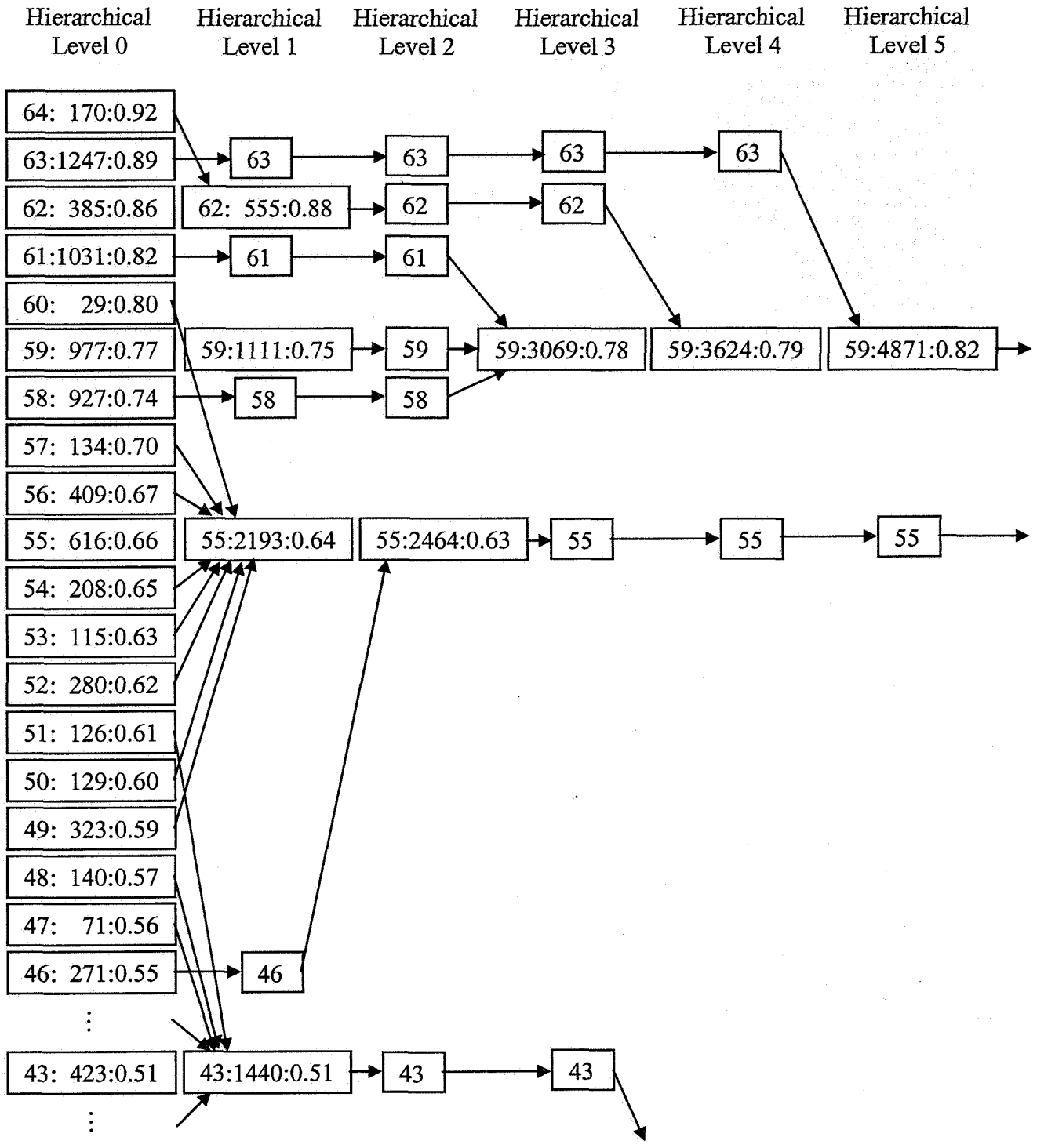
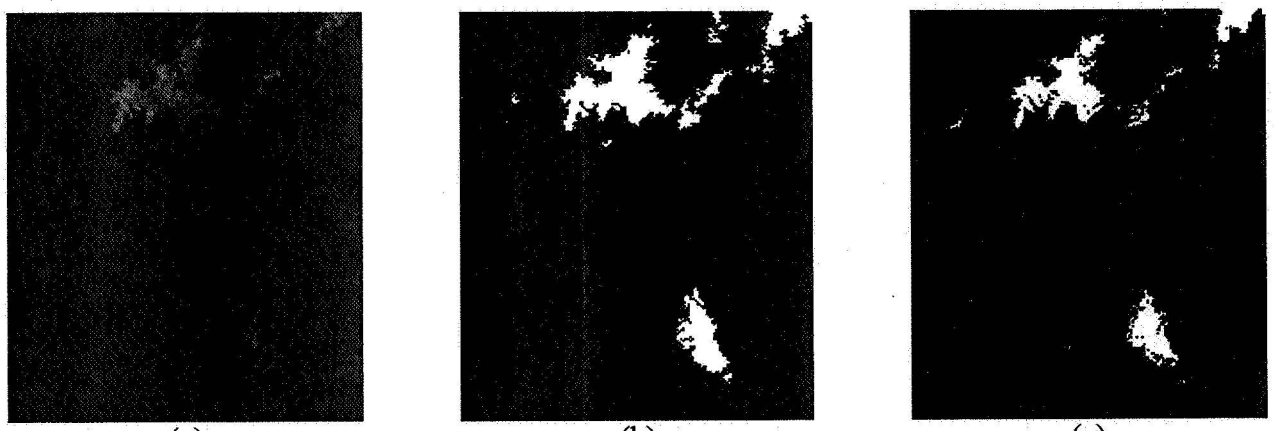


Figure 7. A portion of the segmentation hierarchy for the RHSEG *spclust_wght* = 0.1 results for the snow/mask generation task. At each node in the segmentation hierarchy is labeled with either the region label alone (unchanged), or with the region label, number of pixels and region mean (label:pixels:mean). The arrows indicate the merging hierarchy. It is significant to note that regions do not necessarily merge with the most similar region in mean value. With *spclust_wght* = 0.1, spatial connectivity influences merge behavior.



(a) (b) (c)

Figure 8. A detailed comparison of the snow/ice mask detected by (a) MODIS composite image (bands 6, 4, 2), (b) RHSEG, and (c) the MODIS standard snow/ice product. The yellow area, obtained by selecting the brightest region at the second coarsest hierarchical level, was very similar to the snow/ice mask provided by the MODIS standard product. An exception to this was the snow field in the lower right. Here we need to add RHSEG region 55 at hierarchical level 7 (the white area) to obtain a snow/ice mask similar to the MODIS standard product.