

16. Image Segmentation Algorithms for Land Categorization

James C. Tilton, Selim Aksoy and Yuliya Tarabalka

CONTENTS

- 16.1. Introduction
- 16.2. Spectrally-based Segmentation Approaches
 - 16.2.1. Thresholding-based Algorithms
 - 16.2.2. Clustering-based Algorithms
 - 16.2.3. Support Vector Machines
- 16.3. Spatially-based Segmentation Approaches
 - 16.3.1. Region Growing Algorithms
 - 16.3.1.1. Seeded Region Growing
 - 16.3.1.2. Split and Merge Region Growing
 - 16.3.1.3. Hybrid of Spectral Clustering and Region Growing
 - 16.3.1.4. Dissimilarity Criterion and Specialized Control Logic
 - 16.3.1.5. Selection of a Single Segmentation from a Segmentation Hierarchy
 - 16.3.2. Texture-based Algorithms
 - 16.3.3. Morphological Algorithms
 - 16.3.3.1. Watershed Algorithms
 - 16.3.3.2. Morphological Profiles
 - 16.3.4. Graph-based Algorithms
 - 16.3.4.1. Optimal Spanning Forests
 - 16.3.4.2. Normalized Cuts
 - 16.3.5. MRF-based Algorithms
 - 16.3.5.1. Simulated Annealing and Iterated Conditional Modes
 - 16.3.5.2. Graph-cuts
- 16.4. Image Segmentation Quality Evaluation
- 16.5. Concluding Remarks
- References

16.1. Introduction

Image segmentation is the partitioning of an image into related meaningful sections or regions. Segmentation is a key first step for a number of image analysis approaches. The nature and quality of the image segmentation result is a critical factor in determining the level of performance of these image analysis approaches. It is expected that an appropriately designed image segmentation approach will provide a better understanding of a landscape, and/or

significantly increase the accuracies of a landscape classification. An image can be partitioned in several ways, based on numerous criteria. Whether or not a particular image partitioning is useful depends on the goal of the image analysis application that is fed by the image segmentation result.

The focus of this chapter is on image segmentation algorithms for land categorization. Our image analysis goal will generally be to appropriately partition an image obtained from a remote sensing instrument on-board a high flying aircraft or a satellite circling the earth or other planet. An example of an earth remote sensing application might be to produce 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 types of other man created development. Alternatively, one can segment the land based on climate (e.g., temperature, precipitation) and elevation zones. However, most image segmentation approaches do not directly provide such meaningful labels to image partitions. Instead, most approaches produce image partitions with generic labels such as region 1, region 2, and so on, which need to be converted into meaningful labels by a post-segmentation analysis.

An early survey on image segmentation grouped image segmentation approaches into three categories (Fu and Mui, 1981): (i) characteristic feature thresholding or clustering, (ii) boundary detection, and (iii) region extraction. Another early survey (Haralick and Shapiro, 1985) divides region extraction into several region growing and region split and merge schemes. Both of these surveys note that there is no general theory of image segmentation, most image segmentation approaches are *ad hoc* in nature, and there is no general algorithm that will work well for all images. This is still the case even today.

We start our image segmentation discussion with spectrally-based approaches, corresponding to Fu and Mui's characteristic feature thresholding or clustering category. We include here a description of support vector machines, as a supervised spectral classification approach that has been a popular choice for analyzing multispectral and hyperspectral images (images with several tens or even hundreds of spectral bands). We then go on to describe a number of spatially-based image segmentation approaches that could be appropriate for land categorization applications, generally going from the simpler approaches to the more complicated and more recently developed approaches. Here our emphasis is guided by the prevalence of reported use in land categorization studies. We then take a brief look at various approaches to image segmentation quality evaluation, and include a closer look at a particular empirical discrepancy approach with example quality evaluations for a particular remotely sensed hyperspectral data set and selected image segmentation approaches. We wrap up with some concluding comments and discussion.

16.2. Spectrally-based Segmentation Approaches

The focus of this section is on approaches that are mainly based on analyses of individual pixels. These approaches use an initial labeling of pixels using unsupervised or supervised classification methods, and then try to group neighboring pixels with similar labels using some form of post-processing to produce segmentation results.

16.2.1. Thresholding-based Algorithms

Thresholding has been one of the oldest and most widely used techniques for image segmentation. Thresholding algorithms used for segmentation assume that the pixels that belong to the objects of interest have a property whose values are substantially different from those of

the background, and aim to find a good set of thresholds that partition the histogram of this property into two or more non-overlapping regions (Sezgin and Sankur, 2004). While the spectral channels can be directly used for thresholding, other derived properties of the pixels are also commonly used in the literature. For example, Akcay and Aksoy (2011) used thresholding of the red band to identify buildings with red roofs, Bruzzone and Prieto (2005) performed change detection by thresholding the difference image, Rosin and Hervas (2005) used thresholding of the difference image for determining landslide activity, Aksoy, et al. (2010) applied thresholding to the normalized difference vegetation index (NDVI) for segmenting vegetation areas, and Unsalan and Boyer (2005) combined thresholding of NDVI and a shadow-water index to identify potential building and street pixels in residential regions.

Selection of the threshold values is often done in an ad hoc manner usually when a single property is involved, but optimal values can also be found by employing exhaustive or stochastic search procedures that look for the values that optimize some criteria on the shape or the statistics of the histogram such as minimization of the within-class variance and maximization of the between-class variance (Otsu, 1979). A stochastic search procedure is particularly needed for finding multiple thresholds where an exhaustive search is not computationally feasible due to the combinatorial increase in the number of candidate values. For example, a recent use of multilevel thresholding for the segmentation of Earth observation data is described in (Ghamisi, et al., 2014) where a particle swarm optimization based stochastic search algorithm was used to obtain a multilevel thresholding of each spectral channel independently by maximizing the corresponding between-class variance.

Even when the selected thresholds are obtained by optimizing some well-defined criteria on the distributions of the properties of the pixels, they do not necessarily produce operational

image segmentation results because they suffer from the lack of the use of spatial information as the decisions are independently made on individual pixels. Thus, thresholding is usually applied as a pre-processing algorithm, and various post-processing methods such as morphological operations are often applied to the results of pixel-based thresholding algorithms as discussed in the following section.

16.2.2. Clustering-based Algorithms

The clustering-based approaches to image segmentation aim to make use of the rich literature on data grouping and/or partitioning techniques for pattern recognition (Duda, Hart and Stork, 2001). It is intuitive to pose the image segmentation problem as the clustering of pixels, and thus, pixel-based image analysis techniques in the remote sensing literature have found natural extensions to image segmentation. In the most widely used methodology, first, the spectral feature space is partitioned and the individual pixels are grouped into clusters without regard to their neighbors, and then, a post-processing step is applied to form regions by merging neighboring pixels having the same cluster label by using a connected components labeling algorithm.

The initial clustering stage commonly employs well-known techniques such as k-means (Aksoy and Akcay, 2005), fuzzy c-means (Shankar, 2007), and their probabilistic extension using the Gaussian mixture model estimated via expectation-maximization (Fauvel, et al., 2013). Since no spatial information is used during the clustering procedure, pixels with the same cluster label can either form a single connected spatial region, or can belong to multiple disjoint regions that are assigned different labels by the connected components labeling algorithm. This reduces the significance of the difficulty of the user's *a priori* selection of the number of clusters in many popular clustering algorithms as there is no strict correspondence between the initial number of

clusters and the final number of image regions. However, it still has a high potential of producing an oversegmentation consisting of noisy results with isolated pixels having labels different from those of their neighbors due to the lack of the use of spatial data.

Therefore, a following post-processing step aims to produce a smoother and spatially consistent segmentation by converting the pixel-based clustering results into contiguous regions. A popular approach is to use an additional segmentation result (often also an oversegmentation), and to use a majority voting procedure for spatial regularization by assigning each region in the oversegmentation a single label that is determined according to the most frequent cluster label among the pixels in that region (Fauvel, et al., 2013). An alternative approach is to use an iterative split-and-merge procedure as follows (Aksoy, et al., 2005), (Aksoy, 2006):

- 1) Merge pixels with identical labels to find the initial set of regions and mark these regions as foreground,
- 2) Mark regions with areas smaller than a threshold as background using connected components analysis,
- 3) Use region growing to iteratively assign background pixels to the foreground regions by placing a window at each background pixel and assigning it to the class that occurs the most in its neighborhood.

This procedure corresponds to a spatial smoothing of the clustering results. The resulting regions can be further processed using mathematical morphology operators to automatically divide large regions into more compact sub-regions as follows:

- 1) Find individual regions using connected components analysis for each cluster,
- 2) For all regions, compute the erosion transform and repeat:
 - a) Threshold erosion transform at steps of 3 pixels in every iteration,

- b) Find connected components of the thresholded image,
- c) Select sub-regions that have an area smaller than a threshold,
- d) Dilate these sub-regions to restore the effects of erosion,
- e) Mark these sub-regions in the output image by masking the dilation using the original image,

until no more sub-regions are found,

- 3) Merge the residues of previous iterations to their smallest neighbors.

Even though we focused on producing segmentations using clustering algorithms in this section, similar post-processing techniques for converting the pixel-based decisions into contiguous regions can also be used with the outputs of pixel-based thresholding (Section 16.2.1) and classification (Section 16.2.3) procedures (Aksoy, et al., 2005).

It is also possible to pose clustering, and the corresponding segmentation, as a density estimation problem. A commonly used algorithm that combines clustering with density estimation and segmentation is the mean shift algorithm (Comaniciu and Meer, 2002). Mean shift is based on nonparametric density estimation where the local maxima (i.e., modes) of the density can be assumed to correspond to clusters. The algorithm does not require *a priori* knowledge of the number of clusters in the data, and can identify the locations of the local maxima by a set of iterations. These iterations can be interpreted as the shifting of points toward the modes where convergence is achieved when a point reaches a particular mode. The shifting procedure uses a kernel with a scale parameter that determines the amount of local smoothing performed during density estimation. The application of the mean shift procedure to image segmentation uses a separate kernel for the feature (i.e., spectral) domain and another kernel for the spatial (i.e., pixel) domain. The scale parameter for the spectral domain can be estimated by

maximizing the average likelihood of held-out data. The scale parameter for the spatial domain can be selected according to the amount of compactness or oversegmentation desired in the image, or can be determined by using geospatial statistics (e.g., by using semivariogram-based estimates) (Dongping, et al., 2012). Furthermore, agglomerative clustering of the mode estimates can be used to obtain a multi-scale segmentation.

16.2.3. Support Vector Machines

Output of supervised classification of pixels can also be used as input for segmentation techniques. In recent years, support vector machines (SVM) and the use of kernels to transform data into a new feature space where linear separability can be exploited have been proposed. The SVM method attempts to separate training samples belonging to different classes by tracing maximum margin hyperplanes in the space where the samples are mapped. SVM have shown to be particularly well suited to classify high-dimensional data (e.g. hyperspectral images) when a limited number of training samples is available (Camps-Valls, 2005), (Vapnik, 1998). The success of SVM for pixel-based classification has led to its subsequent use as part of image segmentation methods. Thus we discuss the SVM approach in detail below.

SVM are primarily designed to solve binary tasks, where the class labels take only two values: 1 or -1. Let us consider a binary classification problem in a B -dimensional space \mathbb{R}^B , with N training samples, $x_i \in \mathbb{R}^B$, and their corresponding class labels $y_i = \pm 1$ available. The SVM technique consists in finding the hyperplane that maximizes the margin, i.e., the distance to the closest training data points in both classes (see Figure 16-1). Noting $\mathbf{w} \in \mathbb{R}^B$ as the vector normal to the hyperplane and $b \in \mathbb{R}$ as the bias, the hyperplane H is defined as

$$\mathbf{w} \cdot x + b = 0, \forall x \in H.$$

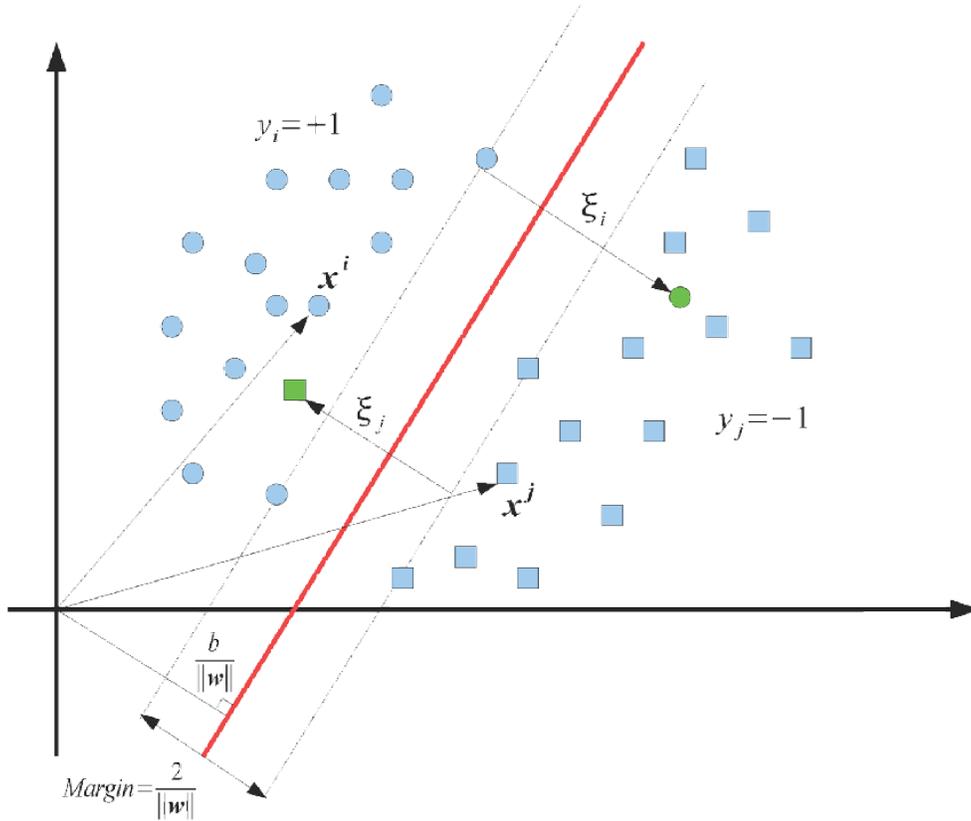


Figure 16-1. Schematic illustration of the SVM binary classification method. There is one non-linearly separable sample in each class.

If $x \notin H$ then

$$f(x) = \frac{|\mathbf{w} \cdot x + b|}{\|\mathbf{w}\|}$$

defines the distance of the sample x to H . In the linearly separable case, such a hyperplane must satisfy:

$$y_i(\mathbf{w} \cdot x_i + b) > 1, \forall i \in [1, N]. \quad (16-1)$$

The optimal hyperplane is the one that maximizes the margin $2/\|\mathbf{w}\|$. This is equivalent to minimizing $\|\mathbf{w}\|/2$ and leads to the quadratic optimization problem:

$$\min \left[\frac{\|\mathbf{w}\|^2}{2} \right], \text{ subject to (16-1)}. \quad (16-2)$$

To take into account non-linearly separable data, slack variables ξ are introduced to deal with misclassified samples (see Figure 16-1). Equation (16-1) becomes

$$y_i(\mathbf{w} \cdot \mathbf{x}_i + b) > 1 - \xi_i, \quad \xi_i \geq 0, \quad \forall i \in [1, N]. \quad (16-3)$$

The final optimization problem is formulated as

$$\min \left[\frac{\|\mathbf{w}\|^2}{2} + C \sum_{i=1}^N \xi_i \right], \text{ subject to (16-3)}. \quad (16-4)$$

where the constant C is a regularization parameter that controls the amount of penalty. This optimization problem is typically solved by quadratic programming (Vapnik, 1998). The classification is further performed by computing $y_u = \text{sign}(\mathbf{w} \cdot \mathbf{x}_u + b)$, where (\mathbf{w}, b) are the hyperplane parameters found during the training process and \mathbf{x}_u is an unseen sample.

One can notice that the pixel vectors in the optimization and decision rule equations always appear in pairs related through a scalar product. These products can be replaced by nonlinear functions of the pairs of vectors, essentially projecting the pixel vectors in a higher dimensional space \mathbb{H} and thus improving linear separability of data:

$$\begin{aligned} \mathbb{R}^B &\rightarrow \mathbb{H}, \\ x &\rightarrow \Phi(x), \end{aligned} \quad (16-5)$$

$$x_i \cdot x_j \rightarrow \Phi(x_i) \cdot \Phi(x_j) = K(x_i, x_j).$$

Here, $\Phi(\cdot)$ is a nonlinear function to project feature vectors into a new space, $K(\cdot)$ is a *kernel* function, which allows one to avoid the computation of scalar products in the transformed space $[\Phi(x_i) \cdot \Phi(x_j)]$ and thus reduces the computational complexity of the algorithm. The

kernel K must satisfy Mercer's condition (Burges, 1998). The Gaussian Radial Basis Function (RBF) kernel is the most widely used for remote sensing image classification:

$$K_{Gaussian}(x_i, x_j) = \exp[-\gamma \|x_i - x_j\|^2], \quad (16-6)$$

where γ is the spread of the RBF kernel.

To solve the K -class problem, various approaches have been proposed. Two main approaches combining a set of binary classifiers are defined as (Smola, 2002):

- One versus all: K binary classifiers are applied on each class against the others. Each pixel x_i is assigned to the class with the maximum output $f(x_i)$.
- One versus one: $K(K - 1)/2$ binary classifiers are applied on each pair of classes. Each pixel is assigned to the class winning the maximum number of binary classification procedures.

As a conclusion, SVM directly exploit the geometrical properties of data, without involving a density estimation procedure. This method has proven to be more effective than other nonparametric classifiers (such as neural networks or the k-Nearest Neighbor classifier (Duda, Hart, & Stork, 2001)) in terms of classification accuracies, computational complexity and robustness to parameter setting. SVM can efficiently handle high-dimensional data, exhibiting low sensitivity to the Hughes phenomenon (Hughes, 1968). Finally, it exhibits good generalization capability, fully exploiting the discrimination capability of available training samples.

Pixel-based supervised classification results, such as those obtained using an SVM classifier, are often given as input to segmentation procedures that aim to group the pixels to form contiguous regions as discussed in the following sections.

16.3. Spatially-based Segmentation Approaches

We cannot possibly discuss myriad of spatially-based image segmentation approaches that have been proposed and developed over the years. Instead we will focus on approaches that have achieved demonstrated success in remote sensing land categorization applications. A compilation of such approaches can be found in a series of papers published by a research group based at the Leibniz Institute for Ecological and Regional Development (IOER) that present comparative evaluations of image segmentation approaches implemented in various image analysis packages (Meinel and Neubert, 2004), (Neubert, et al., 2006), (Neubert, et al., 2008), (Marpu, et al., 2010). Table 16.1 provides a summary listing of most of the remote sensing oriented image analysis packages whose image segmentation approach was evaluated in these papers, plus image segmentation approaches from three additional notable remote sensing oriented image analysis packages (GRASS GIS, IDRISI and the Orfeo toolbox).

We note from Table 16.1 that region growing is the most frequent image segmentation approach utilized by these remote sensing oriented image analysis software packages. Further, several packages combine region growing with other techniques (RHSeg with spectral clustering, SCRM and GRASS GIS with watershed, SegSAR with edge detection). Watershed segmentation (an approach based on region boundary detection) is the next most popular approach. Simulated annealing is often utilized in analysis packages oriented towards analyzing Synthetic Aperture Radar (SAR) imagery data (Ceasar and InfoPACK).

The next several sections describe various spatially-based image segmentation approaches, starting with region growing algorithms, and continuing with texture-based algorithms, morphological algorithms, graph-based algorithms, and MRF-based algorithms.

Table 16.1. The algorithmic basis of image segmentation approaches in remote sensing oriented image analysis packages. Most of these image segmentation approaches were evaluated in a series of papers by the Leibnitz IOER group.

Name	Website or Reference	Algorithmic Basis
BerkeleyImgseg	http://www.imageseg.com/	Region-growing, region-merging
Ceasar	(Cook, et al., 1996)	Simulated annealing
eCognition Developer ¹	http://www.ecognition.com/products/ecognition-developer and (Baatz and Schape, 2000)	Region growing
ENVI Feature Extraction	http://www.exelisvis.com/docs/using_envi_Home.html and (Robinson, et al., 2002)	Edge-based (Full Lambda-Schedule algorithm for region merging)
Extended Watershed EWS	(Li and Xiao, 2007)	Multi-channel watershed transformation
Image WS for Erdas Imagine	(Sramek and Wrbka, 1997)	Hierarchical watershed
InfoPACK	(Cook, et al., 1996)	Simulated annealing
PARBAT	http://parbat.lucieer.net/ and (Lucieer, 2004)	Region growing
RHSeg	(Tilton, 1998) and (Tilton, et al., 2012)	Region growing and spectral clustering
SCRM	(Castilla, et al., 2008)	Watershed and region merging
SegSAR	http://www.segsar.googlepages.com/	Hybrid (edge/region oriented)
SEGEN	(Gofman, 2006)	Region growing
GRASS GIS	http://grasswiki.osgeo.org/wiki/GRASS-Wiki	Region growing and watershed
IDRISI	http://www.clarklabs.org/applications/upload/Segmentation-IDRISI-Focus-Paper.pdf	Watershed
Orfeo Toolbox	http://www.orfeo-toolbox.org/otb/	Region growing, watershed, level sets, mean-shift

Note: 1. Was Definiens Developer – but the remote sensing package is now marketed by Trimble, and the Definiens product is now oriented to biomedical image analysis.

16.3.1. Region Growing Algorithms

In the region growing approach to image segmentation an image is initially partitioned into small region objects. These initial small region objects are often single image pixels, but can also be $n \times n$ blocks of pixels or another partitioning of the image into small spatially connected region objects. Then pairs of spatially adjacent region objects are compared and merged together if they are found to be similar enough according to some comparison criterion. The underlying assumption is that region objects of interest are several image pixels in size and relatively homogeneous in value. Most region growing approaches can operate on either grey scale, multispectral or hyperspectral image data, depending on the criterion used to determine the similarity between neighboring region objects.

A very early example of region growing was described in (Muerle and Allen, 1968). Muerle and Allen experimented with initializing their region growing process with region objects consisting of 2×2 up to 8×8 blocks of pixels. After initialization, they started with the region object at the upper left corner of the image and compared this region object with the neighboring region objects. If a neighboring region object was found to be similar enough, the region objects were merged together. This process was continued until no neighboring region objects could be found that were similar enough to be merged into the region object that was being grown. Then the image was scanned (left-to-right, top-to-bottom) to find an unprocessed region object, *i.e.*, a region object that not yet been considered as an initial object for region growing or merged into a neighboring region. If an unprocessed region object was found, they conducted their region growing process from that region object. This continued until no further unprocessed region objects could be found, upon which point the region growing segmentation process was considered completed.

Many early schemes for region growing, such as Muerle and Allen's, can be formulated as logical predicate segmentation, defined in (Zucker, 1976) as:

A segmentation of an image X can be defined as a partition of X into R disjoint subsets X_1, X_2, \dots, X_R , such that the following conditions hold:

- 1) $\bigcup_{i=1}^R X_i = X$,
- 2) $X_i, i = 1, 2, \dots, R$ are connected,
- 3) $P(X_i) = \text{TRUE}$ for $i = 1, 2, \dots, R$, and
- 4) $P(X_i \cup X_j) = \text{FALSE}$ for $i \neq j$, where X_i and X_j are adjacent.

$P(X_i)$ is a logical predicate that assigns the value TRUE or FALSE to X_i , depending on the image data values in X_i .

These conditions are summarized in (Zucker, 1976) as follows: The first condition requires that every picture element (pixel) must be in a region (subset). The second condition requires that each region must be connected, *i.e.* composed of contiguous image pixels. The third condition determines what kind of properties each region must satisfy, *i.e.* what properties the image pixels must satisfy to be considered similar enough to be in the same region. Finally, the fourth condition specifies that any merging of adjacent regions would violate the third condition in the final segmentation result.

Several researchers in this early era of image segmentation research, including Muerle and Allen, noted some problems with logical predicate segmentation. For one, the results were very dependent on the order in which the image data was scanned. Also, the statistics of a region object can change quite dramatically as the region is grown, making it possible that many adjacent cells that were rejected for merging early in the region growing process would have been accepted in later stages based on the changed statistics of the region object. The reverse was

also possible, where adjacent cells that were accepted for merging early in the region growing process would have been rejected in later stages.

Subsequently, an alternate approach to region growing was developed that avoids the above mentioned problems, and approach that eventually became to be referred to as best merge region growing. An early version of best merge region growing, hierarchical step-wise optimization (HSWO), is an iterative form of region growing, in which the iterations consist of finding the most optimal or best segmentation with one region less than the current segmentation (Beaulieu and Goldberg, 1989). The HSWO approach can be summarized as follows:

- i. 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.
- ii. Calculate the dissimilarity criterion value, d , between all pairs of spatially adjacent regions, find the smallest dissimilarity criterion value, T_{merge} , and merge all pairs of regions with $d = T_{merge}$.
- iii. Stop if no more merges are required. Otherwise, return to step ii.

HSWO naturally produces a segmentation hierarchy consisting of the entire sequence of segmentations from initialization down to the final trivial one region segmentation (if allowed to proceed that far). For practical applications, however, a subset of segmentations needs to be selected out from this exhaustive segmentation hierarchy. At a minimum such a subset can be defined by storing the results only after a preselected number of regions is reached, and then storing selecting iterations after that such that no region is involved in more than one merge between stored iteration until a two region segmentation is reached. A portion of such a

segmentation hierarchy is illustrated in Figure 16-2. (The selection of a single segmentation from a segmentation hierarchy is discussed in Section 16.3.1.5.)

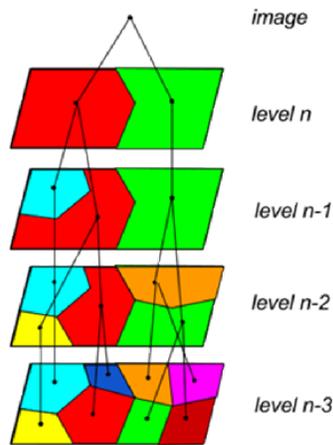


Figure 16-2. The last four levels of an n level segmentation hierarchy produced by a region growing segmentation process. Note that when depicted in this manner, the region growing process is a “bottom-up” approach.

A unique feature of the segmentation hierarchy produced by HSWO and related region growing segmentation approaches is that the segment or region boundaries are maintained at the full image spatial resolution for all levels of the segmentation hierarchy. The region boundaries are coarsened in many other multilevel representations.

Many variations on best merge region growing have been described in the literature. As early as 1994, Kurita (1994) described an implementation of the HSWO form of best merge region growing that utilized a heap data structure (Williams, 1964) for efficient determination of best merges and a dissimilarity criterion based on minimizing the mean squared error between the region mean image and original image. Further, several of the image segmentation approaches studied in the previously referenced series of papers published by the Leibniz Institute are based on best merge region growing (see Table 16-1). As we will discuss in more detail later, the main differences between most of these region growing approaches are the

dissimilarity criterion employed and, perhaps, some control logic designed to remove small regions or otherwise tailor the segmentation output.

16.3.1.1. Seeded Region Growing

Seeded region growing is a variant on best merge region growing in which regions are grown from pre-selected seed pixels or seed regions. Adams and Bischof (1994) is an early example of this approach. In seeded region growing, the best merges are found by examining the pixels adjacent to each seed pixel of region formed around each seed pixel. As described by Adams and Bischof (1994), the region growing process continues until each image pixel is associated with one of the pre-selected seed pixels or regions.

16.3.1.2. Split and Merge Region Growing

In the split and merge approach, the image is repeatedly subdivided until each resulting region has a sufficiently high homogeneity. Examples of measures of homogeneity are the mean squared error between the region mean and the image data values, or the region standard deviation. After the region-splitting process converges, the regions are grown using one of the previously described region growing approaches. This approach is more efficient when large homogeneous regions are present. However, some segmentation detail may be lost. See (Horowitz and Pavlidis, 1974), (Cross, et al., 1988), and (Strasters and Gerbrands, 1991) for examples of this approach.

16.3.1.3. Hybrid of Spectral Clustering and Region Growing

Tilton (1998) and Tilton, et al. (2012) describe a hybridization of HSWO best merge region growing with spectral clustering, called HSeg (for Hierarchical Segmentation). We remind the

reader that HSWO is performed by finding a threshold value, T_{merge} , equal to the value of a dissimilarity criterion of the most similar pair of spatially adjacent regions, and then merging all pairs of regions that have dissimilarity equal to T_{merge} . HSeg adds to HSWO a step following each step of adjacent region merges in which all pairs of spatially non-adjacent regions are merged that have dissimilarity $\leq S_w T_{merge}$, where $0.0 \leq S_w \leq 1.0$ is a factor that sets the priority between spatially adjacent and non-adjacent region merges. Note that when $S_w = 0.0$, HSeg reduces to HSWO.

Unfortunately, the inclusion of the step in HSeg of merging spatially non-adjacent regions adds significantly to the computational requirements of this image segmentation approach. This is because comparisons must now be made between all pairs of regions instead of just between pairs of spatially adjacent regions. A recursive divide-and-conquer approach (called RHSeg), along with a parallel implementation, was developed to help overcome this problem (Tilton, 2007). The computational requirements of this approach were further reduced by a refinement in which the non-adjacent region merging was limited to between regions of a minimum size, P_{min} (Tilton, et al., 2012). In this refinement the value of P_{min} is dynamically adjusted to keep the number of “large regions” (those of at least P_{min} in size) to a range that substantially reduces the computational requirements without out significantly affecting the image segmentation results.

16.3.1.4. Dissimilarity Criterion and Specialized Control Logic

Muerle and Allen (1968) experimented with various criteria for determining whether or not pairs of region objects were similar enough to be merged together. They concluded that an optimal criterion would be a threshold of a function of the mean and standard deviation of the gray levels of the pixels contained in the compared region objects.

In our image segmentation research, we have implemented and studied several region merging criteria in the form of dissimilarity criteria (Tilton, 2013). Included among these dissimilarity criteria are criteria based on vector norms (the 1-, 2-, and ∞ -norms), criteria based on minimizing the increase of mean squared error between the region mean image and the original image data, and a criterion based on the Spectral Angle Mapper (SAM) criterion (Kruse, et al., 1993). We briefly describe these dissimilarity criteria here.

The dissimilarity criterion based on the 1-Norm of the difference between the region mean vectors, u_i and u_j , of regions X_i and X_j , each with B spectral bands, is:

$$d_{1\text{-Norm}}(X_i, X_j) = \|u_i - u_j\|_1 = \sum_{b=1}^B |\mu_{ib} - \mu_{jb}|, \quad (16-7)$$

where μ_{ib} and μ_{jb} are the mean values for regions i and j , respectively, in spectral band b , i.e., $u_i = (\mu_{i1}, \mu_{i2}, \dots, \mu_{iB})^T$ and $u_j = (\mu_{j1}, \mu_{j2}, \dots, \mu_{jB})^T$.

The dissimilarity criterion based on the 2-Norm is:

$$d_{2\text{-Norm}}(X_i, X_j) = \|u_i - u_j\|_2 = \left[\sum_{b=1}^B (\mu_{ib} - \mu_{jb})^2 \right]^{1/2}, \quad (16-8)$$

and the dissimilarity criterion based on the ∞ -Norm is:

$$d_{\infty\text{-Norm}}(X_i, X_j) = \|u_i - u_j\|_{\infty} = \max(|\mu_{ib} - \mu_{jb}|, b = 1, 2, \dots, B) \quad (16-9)$$

As noted above, a criterion based on mean squared error minimizes the increase of mean squared error between the region mean image and the original image data as regions are grown.

The sample estimate of the mean squared error for the segmentation of band b of the image X into R disjoint subsets X_1, X_2, \dots, X_R is given by:

$$MSE_b(X) = \frac{1}{N-1} \sum_{i=1}^R MSE_b(X_i), \quad (16-10a)$$

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

$$MSE_b(X_i) = \sum_{x_p \in X_i} (\chi_{pb} - \mu_{ib})^2 \quad (16-10b)$$

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

$$d_{BSMSE}(X_i, X_j) = \sum_{b=1}^B \Delta MSE_b(X_i, X_j) \quad (16-11a)$$

where

$$\Delta MSE_b(X_i, X_j) = MSE_b(X_i \cup X_j) - MSE_b(X_i) - MSE_b(X_j). \quad (16-11b)$$

$BSMSE$ refers to “band sum MSE .” Using (16-10b) and exchanging the order of summation, (16-11b) can be manipulated to produce an efficient dissimilarity function based on aggregated region features (for the details, see (Tilton, 2013)):

$$d_{BSMSE}(X_i, X_j) = \frac{n_i n_j}{(n_i + n_j)} \sum_{b=1}^B (\mu_{ib} - \mu_{jb})^2. \quad (16-12)$$

The dimensionality of the d_{BSMSE} dissimilarity criteria 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 dimensionalities consistent, the square root of d_{BSMSE} is often used.

The Spectral Angle Mapper (SAM) criterion is widely used in hyperspectral image analysis (Kruse, et al., 1993). This criterion determines the spectral similarity between two spectral vectors by calculating the “angle” between the two spectral vectors. An important property of the SAM criterion is that poorly illuminated and more brightly illuminated pixels of the same color

will be mapped to the same spectral angle despite the difference in illumination. The spectral angle θ between the region mean vectors, u_i and u_j , of regions X_i and X_j is given by:

$$\theta(u_i, u_j) = \arccos\left(\frac{u_i \circ u_j}{\|u_i\|_2 \|u_j\|_2}\right) = \arccos\left(\frac{\sum_{b=1}^B \mu_{ib} \mu_{jb}}{\left(\sum_{b=1}^B \mu_{ib}^2\right)^{1/2} \left(\sum_{b=1}^B \mu_{jb}^2\right)^{1/2}}\right) \quad (16-13a)$$

where μ_{ib} and μ_{jb} are the mean values for regions i and j , respectively, in spectral band b , i.e., $u_i = (\mu_{i1}, \mu_{i2}, \dots, \mu_{iB})^T$ and $u_j = (\mu_{j1}, \mu_{j2}, \dots, \mu_{jB})^T$. The dissimilarity function for regions X_i and X_j , based on the SAM distance vector measure, is given by:

$$d_{SAM}(X_i, X_j) = \theta(u_i, u_j) \quad (16-13b)$$

Note that the value of d_{SAM} ranges from 0.0 for similar vectors up to $\pi/2$ for the most dissimilar vectors.

A problem that can often occur with basic best merge region growing approaches is that the segmentation results contain many small regions. We have found this to be the case when employing dissimilarity criteria based on vector norms or SAM, but not a problem for dissimilarity criteria based on minimizing the increase of mean squared error. This is because the mean squared error criterion has a factor, $n_i n_j / (n_i - n_j)$, where n_i and n_j are the number of pixels in the two compared regions, that biases toward merging small regions into larger ones. We have found it useful to add on a similar “small region merge acceleration factor” to the vector norm and SAM based criterion when one of the compared regions is smaller than a certain size. See Tilton, et al. (2012) for more details.

Implementations of best merge region growing often add special control logic to reduce the number of small regions or otherwise improve the final classification result. An example of this is SEGEN (Gofman, 2006), which uses the vector 2-norm (otherwise known as Euclidean distance) for the dissimilarity criterion. As noted in Tilton, et al. (2012), SEGEN is a relatively pure implementation of best merge region growing, optimized for efficiency in performance, memory utilization, and image segmentation quality. SEGEN adds a number of (optional) procedures to best merge region growing, among them a low-pass filter to be applied on the first stage of the segmentation and outlier dispatching on the last stage. The latter removes outlier pixels and small segments by imbedding them in neighborhood segments with the smallest dissimilarity. SEGEN also provides several parameters to control the segmentation process. A set of “good in average” control values is suggested in (Gofman, 2006).

The best merge region growing segmentation approach employed in eCognition Developer utilizes a dissimilarity function that balances minimizing the increase of heterogeneity, f , in both color and shape (Baatz and Schape, 2000), (Benz, et al., 2004):

$$f = w_{color} \cdot \Delta h_{color} + w_{shape} \cdot \Delta h_{shape} \quad (16-14)$$

where w_{color} and w_{shape} are weights that range in value from 0 to 1 and must mutually sum up to 1.

The color component of heterogeneity, Δh_{color} , is defined as follows:

$$\Delta h_{color} = \sum_c w_c \left(n_{merge} \cdot \sigma_{c,merge} - (n_1 \cdot \sigma_{c,1} + n_2 \cdot \sigma_{c,2}) \right) \quad (16-15)$$

where n_x is the number of pixels, with the subscript $x=merge$ referring to the merged object, and subscripts $x = 1$ or 2 referring to the first and second objects considered for merging. σ_c refers to the standard deviation in channel (spectral band) c , with the same additional subscripting denoting the merged or pair of considered objects. w_c is a channel weighting factor.

The shape component of heterogeneity, Δh_{shape} , is defined as follows:

$$\Delta h_{shape} = w_{compt} \cdot \Delta h_{compt} + w_{smooth} \cdot \Delta h_{smooth} \quad (16-16)$$

where

$$\Delta h_{compt} = n_{merge} \cdot \frac{l_{merge}}{\sqrt{n_{merge}}} - \left(n_1 \cdot \frac{l_1}{\sqrt{n_1}} + n_2 \cdot \frac{l_2}{\sqrt{n_2}} \right) \quad (16-17)$$

$$\Delta h_{smooth} = n_{merge} \cdot \frac{l_{merge}}{b_{merge}} - \left(n_1 \cdot \frac{l_1}{b_1} + n_2 \cdot \frac{l_2}{b_2} \right) \quad (16-18)$$

where l is the perimeter of the object, and b is the perimeter of the object's bounding box. The weights w_c , w_{color} , w_{shape} , w_{smooth} and w_{compt} can be selected to best suit a particular application.

The level of detail, or scale, of the segmentation is set by stopping the best merge region growing process when the increase in heterogeneity, f , for the best merge reaches a predefined threshold value. A multiresolution segmentation can be created by performing this process with a set of increasing thresholds.

In eCognition Developer, other segmentation procedures can be combined with the multiresolution approach. For example, spectral difference segmentation merges neighbor objects that fall within a user-defined maximum spectral difference. This procedure can be used to merge spectrally similar objects from the segmentation produced by the multiresolution approach.

16.3.1.5. Selection of a Single Segmentation from a Segmentation Hierarchy

Some best merge region growing approaches such as HSWO and HSeg do not produce a single segmentation result, but instead produce a segmentation hierarchy. The best merge region growing segmentation approach employed in eCognition Developer can also produce a segmentation hierarchy. 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 a 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 by tracking the behavior of these characteristics throughout different levels of detail. A manual approach for doing this using a graphical user interface is described in (Tilton, 2003) and (Tilton, 2013). A preliminary study on automating this approach is described in (Plaza and Tilton, 2005) where it was proposed to automate this approach using joint spectral/spatial homogeneity scores computed from segmented regions. An alternate approach for making spatially localized selections of segmentation detail based on matching region boundaries with edges produced by an edge detector is described in (Le Moigne and Tilton, 1995).

Tarabalka, et al. (2012) proposed a modification of HSeg through which a single segmentation output is automatically selected for output from the usual segmentation hierarchy. The idea is similar to the previously described seeded region growing. The main idea behind the marker-based HSeg algorithm consists in automatically selecting seeds, or *markers*, for image regions, and then performing best merge region growing with an additional condition: two regions with different marker labels cannot be merged together (see Figure 16-3). The authors proposed to choose markers of spatial regions by analyzing results of a probabilistic supervised classification of each pixel and by retaining the most reliably classified pixels as region seeds.

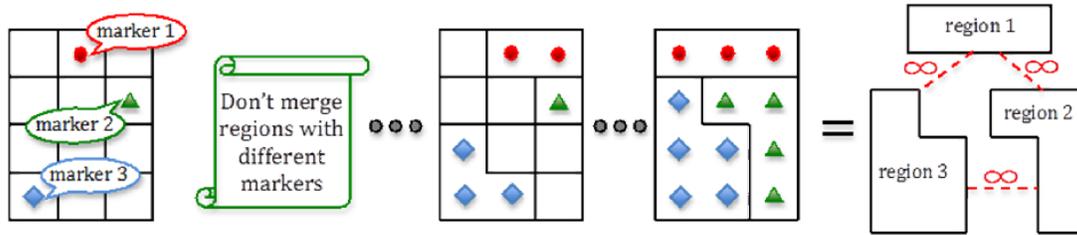


Figure 16-3. Scheme illustrating the marker-based HSeg algorithm.

An alternative algorithm that produces a final segmentation by automatically selecting subsets of regions appearing in different levels in the segmentation hierarchy is described in Section 16.3.4.2

16.3.2. Texture-based Algorithms

Along with spectral information, textural features have also been heavily used for various image analysis tasks, including their use as features in thresholding-based, clustering-based, and region growing-based segmentation in the literature. In this section, we focus on the use of texture for *unsupervised* image segmentation. In the following, first, we discuss some particular examples that involve texture modeling for segmenting natural landscapes and man-made structures using local image properties, and then, present recent work on generalized texture algorithms that aim to model complex image structures in terms of the statistics and spatial arrangements of simpler image primitives.

One of the common uses of texture in remote sensing image analysis is the identification of natural landscapes. For example, Epifanio and Soille (2007) used morphological transformations such as white and black top-hat by reconstruction and thresholded image gradients with an unsupervised clustering procedure where the texture prototypes were automatically selected based on the dissimilarities between the feature vectors of neighboring image windows to segment vegetation zones and forest stands, and Wang and Boesch (2007) combined an initial

color-based oversegmentation step with a threshold-based region merging procedure that used wavelet feature statistics inside candidate image regions to delineate forest boundaries.

Man-made structures also exhibit particular characteristics that can be modeled using textural features. For example, Pesaresi, et al. (2008) proposed a rotation invariant anisotropic texture model that used contrast features computed from gray level co-occurrence matrices, and used this model to produce a built-up presence index. The motivation behind the use of the contrast-based features was to exploit the relationships between the buildings and their shadows. Sirmacek and Unsalan (2010) and Gueguen, et al. (2012) used a similar idea and modeled urban areas using spatial voting (smoothing) of local feature points extracted from the local maxima of Gabor filtering results and corner detection results, respectively. All of these results can be converted to a segmentation output by thresholding the corresponding urban area estimates.

An open problem in image segmentation is to identify the boundaries of regions that are not necessarily homogeneous with respect to low-level features such as color and texture that are extracted from individual pixels or from small local pixel neighborhoods. Recent literature includes several examples that can be considered as generalized texture algorithms that aim to model heterogeneous image content in terms of the spatial arrangements of relatively homogeneous image primitives. Some of this work has considered the segmentation of particular structures that have specific textural properties. For example, Dogrusoz and Aksoy (2007) aimed the segmentation of regular and irregular urban structures by modeling the image content using a graph where building objects and the Voronoi tessellation of their locations formed the vertices and the edges, respectively, and the graph was clustered by thresholding its minimum spanning tree so that organized (formal) and unorganized (informal) settlement patterns were extracted to model urban development. Some agricultural structures such as permanent crops also exhibit

specific textural properties that can be useful for segmentation. For example, Aksoy, et al. (2012) proposed a texture model that is based on the idea that textures are made up of primitives (trees) appearing in a near-regular repetitive arrangement (planting patterns), and used this model to compute a regularity score for different scales and orientations by using projection profiles of multi-scale isotropic filter responses at multiple orientations. Then, they illustrated the use of this model for segmenting orchards by iteratively merging neighboring pixels that have similar regularity scores at similar scales and orientations.

More generic approaches have also been proposed for segmenting heterogeneous structures. For example, Gaetano, et al. (2009) started with an oversegmentation of atomic image regions, and then performed hierarchical texture segmentation by assuming that frequent neighboring regions are strongly related. These relations were represented using Markov chain models computed from quantized region labels, and the image regions that exhibit similar transition probabilities were clustered to construct a hierarchical set of segmentations. Zamalieva, et al. (2009) used a similar frequency-based approach by finding the significant relations between neighboring regions as the modes of a probability distribution estimated using the continuous features of region co-occurrences. Then, the resulting modes were used to construct the edges of a graph where a graph mining algorithm was used to find subgraphs that may correspond to atomic texture primitives that form the heterogeneous structures. The final segmentation was obtained by using the histograms of these subgraphs inside sliding windows centered at individual pixels, and by clustering the pixels according to these histograms. As an alternative to graph-based grouping, Akcay, et al. (2010) performed Gaussian mixture-based clustering of the region co-occurrence features to identify frequent region pairs that are merged in each iteration of a hierarchical texture segmentation procedure.

In addition to using co-occurrence properties of neighboring regions to exploit statistical information, structural features can also be extracted to represent the spatial layout for texture modeling. For example, Akcay and Aksoy (2011) described a procedure for finding groups of aligned objects by performing a depth-first search on a graph representation of neighboring primitive objects. After the search procedure identified aligned groups of three or more objects that have centroids lying on a straight line with uniform spacing, an agglomerative hierarchical clustering algorithm was used to find larger groups of primitive objects that have similar spatial layouts. The approach was illustrated in the finding of groups of buildings that have different statistical and spatial characteristics that cannot be modeled using traditional segmentation methods. Another approach for modeling urban patterns using hierarchical segmentations extracted from multiple images of the same scene at various resolutions was described by (Kurtz, et al., 2012) where binary partition trees were used to model image data and tree cuts were learned from user-defined segmentation examples for interactive partitioning of images into semantic heterogeneous regions.

16.3.3. Morphological Algorithms

Mathematical morphology has been successfully used for various tasks such as image filtering for smoothing or enhancement, texture analysis, feature extraction, and detecting objects with certain shapes in the remote sensing literature (Soille and Pesaresi, 2002), (Soille, 2003). Morphological algorithms have also been one of the most widely used techniques for segmenting remotely sensed images. These approaches view the two-dimensional image data that consist of the spectral channels or some other property of the pixels as an imaginary topographic relief where higher pixel values map to higher imaginary elevation levels (see Figure 16-A). Consequently, differences in the elevations of the pixels in a spatial neighborhood can be

exploited to partition those pixels into different regions. Two morphological approaches for segmentation have found common use in the literature: watershed algorithms and morphological profiles. These approaches are described in the following sections. Other approaches using mathematical morphology for image segmentation, and particularly for producing segmentation hierarchies, can be found in (Soille, 2008), (Soille and Najman, 2010), (Ouzounis, et al., 2012), and (Perret, et al., 2012).

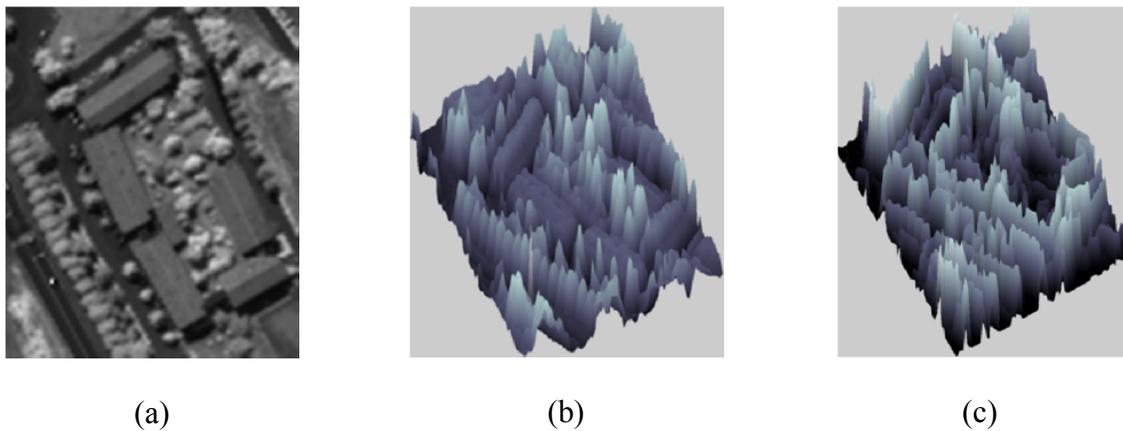


Figure 16-A. Illustration of mapping of the two-dimensional image data that consist of the spectral channels or some other property of the pixels as an imaginary topographic relief so that higher pixel values map to higher imaginary elevation levels. (a) An example spectral band. (b) The spectral values viewed as a three-dimensional topographic relief. (c) Gradient of the spectral data at each pixel viewed as a three-dimensional topographic relief.

16.3.3.1 Watershed Algorithms

The watershed algorithm divides the imaginary topographic relief into catchment basins so that each basin is associated with one local minimum in the image (i.e., individual segments) and the watershed lines correspond to the pixel locations that separate the catchment basins (i.e., segment boundaries). Watershed segmentation can be simulated by an immersion process

(Vincent and Soille, 1991). If we immerse the topographic surface in water, the water rises through the holes at the regional minima with a uniform rate. When two volumes of water coming from two different minima are about to merge, a dam is built at each point of contact. Following the immersion process, the union of all those dams constitutes the watershed lines. A graph-theoretical interpretation of the watershed algorithm can be found in (Meyer, 2001). Couprie, et al. (2011) describes a common framework that unifies watershed segmentation and some other graph-based segmentation algorithms that are described in Section 16.3.4.

The most commonly used method for constructing the topographic relief from the image data to be segmented is to use the gradient function at each pixel. This approach incorporates edge information in the segmentation process, and maps homogeneous image regions with low gradient values into the catchment basins and the pixels in high-contrast neighborhoods with high gradient values into the peaks in the elevation function. The gradient function for single-channel images can easily be computed using derivative filters. Multivariate extensions of the gradient function can be used to apply watershed segmentation to multispectral and hyperspectral images (Aptoula and Lefevre, 2007), (Noyel, et al., 2007), (Li and Xiao, 2007), and (Fauvel, et al., 2013).

A potential problem in the application of watershed segmentation to images with high levels of detail is oversegmentation when the watersheds are computed from raw image gradient where an individual segment is produced for each local minimum of the topographic relief. Pre-processing or post-processing methods can be used to reduce oversegmentation. For example, smoothing filters such as the mean or median filters can be applied to the original image data as a pre-processing step. Alternatively, the oversegmentation produced by the watershed algorithm can be given as input to a region merging procedure for post-processing (Haris, et al., 1998).

Another commonly used alternative to reduce the oversegmentation is to use the concept of dynamics that are related to the regional minima of the image gradient. A regional minimum is composed of a group of neighboring pixels with the same value where the pixels on the external boundary of this group have a greater value. When we consider the image gradient as a topographic surface, the dynamic of a regional minimum can be defined as the minimum height that a point in the minimum has to climb to reach a lower regional minimum (Najman and Schmitt, 1996). The h -minima transform can be used to suppress the regional minima with dynamics less than or equal to a particular value h by performing geodesic reconstruction by erosion of the input image f from $f + h$ (Soille, 2003). When it is difficult to select a single h value, it is common to create a multi-scale segmentation by using an increasing sequence of h values. The multi-scale watershed segmentation generates a set of nested partitions where the partition at scale s is obtained as the watershed segmentation of the image gradient whose regional minima with dynamics less than or equal to s are eliminated by using the h -minima transform. First, the initial partition is calculated as the classical watershed corresponding to all local minima. Next, the two catchment basins having a dynamic of 1 are merged with their neighbor catchment basins at scale 1. Then, at each scale s , the minima with dynamics less than or equal to s are filtered whereas the minima with dynamics greater than s remain the same or are extended. This continues until the last scale corresponding to the largest dynamic in the gradient image. Figure 16-B illustrates the use of the h -minima transform for suppressing regional minima for obtaining a multi-scale watershed segmentation.

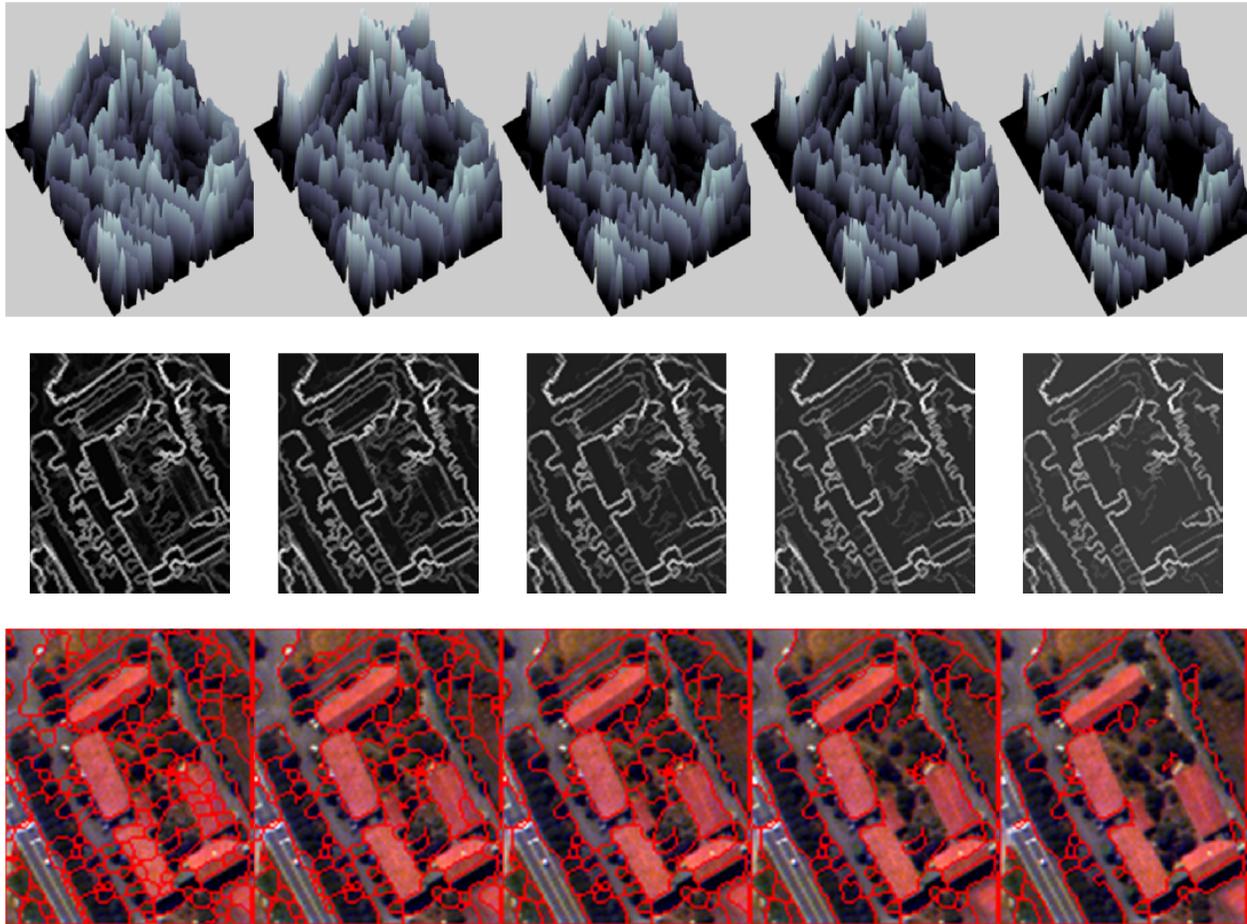


Figure 16-B. Illustration of the h -minima transform for suppressing regional minima for obtaining a multi-scale watershed segmentation. The columns represent increasing values of h , corresponding to decreasing amount of detail in the gradient data. The first row shows the gradient information at each pixel as a topographic relief. The second row shows the gradient data as an image. Brighter values represent higher gradient. The third row shows the segmentation boundaries obtained by the watershed algorithm in red.

Yet another popular approach for computing the watershed segmentation without a significant amount of oversegmentation is to use markers (Meyer and Beucher, 1990). Marker controlled watershed segmentation can be defined as the watershed of an input image

transformed to have regional minima only at the marker locations. Possible methods for identifying the markers include manual selection or selection of the pixels with high confidence values at the end of pixel-based supervised classification (Tarabalka, et al., 2010a). Given a marker image f_m that consists of pixels whose value is 0 at the marker locations and a very large value in the rest of the image, the minima in the input image f can be rearranged by using minima imposition. First, minima can be created only at the locations of the markers by taking the point-wise minimum between $f + 1$ and f_m . Note that the resulting image is lower than or equal to the marker image. The second step of the minima imposition is the morphological reconstruction by erosion of the resulting image from the marker image f_m . Finally, watershed segmentation is applied to the resulting image. It is also possible to produce a multi-scale segmentation by applying marker controlled watershed segmentation to the input image by using a decreasing set of markers. Marker selection is also discussed in Section 16.3.4.

16.3.3.2. Morphological Profiles

The image representation called morphological profiles was popularized in the remote sensing literature by Pesaresi and Benediktsson (2001). The representation uses the morphological residuals between the original image function and the composition of a granulometry constructed at multiple scales. The proposed approach makes use of both classical morphological operators such as opening and closing, and recent theoretical advances such as leveling and morphological spectrum to build the morphological profile.

The fundamental operators in mathematical morphology are erosion and dilation (Soille, 2003). Both of these operators use the definition of a pixel neighborhood with a particular shape called a structuring element (SE) (e.g., a disk of radius of 3 pixels). The erosion operator can be used to identify the image locations where the SE fits the objects in the image, and is defined as

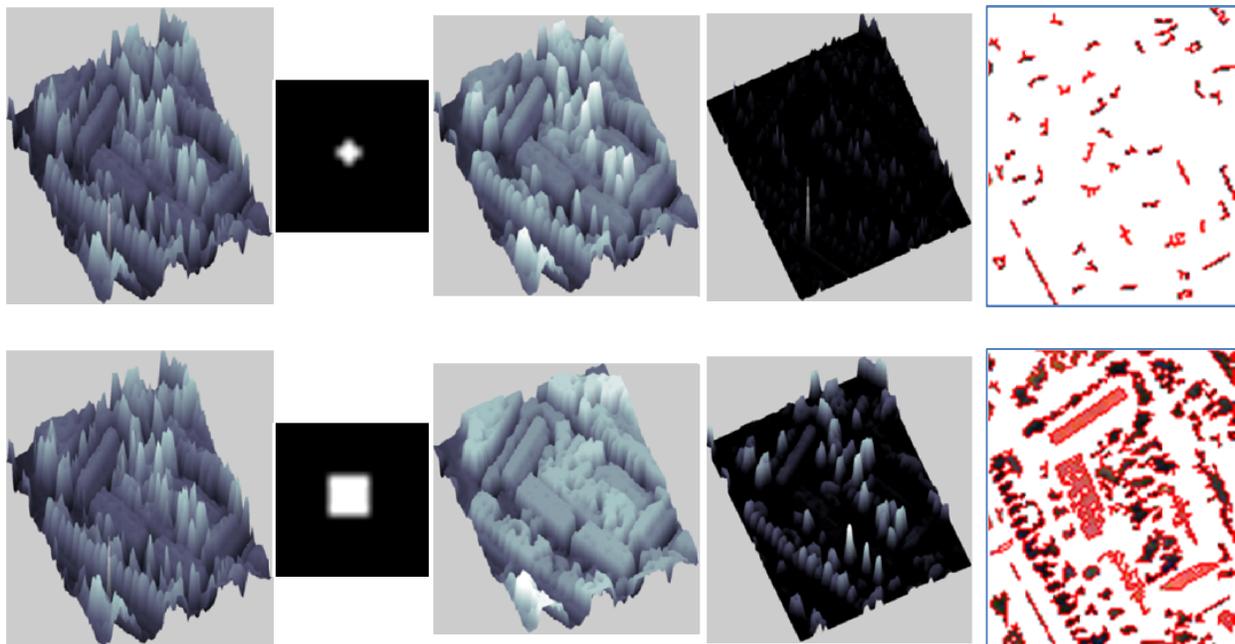
the infimum of the values of the image function in the neighborhood defined by SE. The dilation operator can be used to identify the pixels where the SE hits the objects in the image, and is defined as the supremum of the image values in the neighborhood defined by SE. These two operators can be combined to define other operators. For example, the opening operator, which is defined as the result of erosion followed by dilation using the same SE, can be used to cut the peaks of the topographic relief that are smaller than the SE. On the other hand, the closing operator, which is defined as the result of dilation followed by erosion using the same SE, can be used to fill the valleys that are smaller than the SE.

The morphological operations are often used with the non-Euclidean geodesic metric instead of the classical Euclidean metric (Pesaresi and Benediktsson, 2001). The elementary geodesic dilation of f (called the marker) under g (called the mask) based on SE is the infimum of the elementary dilation of f (with SE) and g . Similarly, the elementary geodesic erosion of f under g based on SE is the supremum of the elementary erosion of f (with SE) and g . A geodesic dilation (respectively, erosion) of size k can also be obtained by performing k successive elementary geodesic dilations (respectively, erosions). Next, the reconstruction by dilation (respectively, erosion) of f under g is obtained by the iterative use of an elementary geodesic dilation (respectively, erosion) of f under g until idempotence is achieved. Then, the opening by reconstruction of an image f can be defined as the reconstruction by dilation of the erosion under the original image. Similarly, the closing by reconstruction of the image f can be defined as the dual reconstruction by erosion of the dilation above the original image.

The advantage of the reconstruction filters is that they do not introduce discontinuities, and therefore, preserve the shapes observed in the input images. Hence, the opening and closing by reconstruction operators can be used to identify the sizes and shapes of different objects present

in the image such that opening (respectively, closing) by reconstruction preserves the shapes of the structures that are not removed by erosion (respectively, dilation), and the residual between the original image and the result of opening (respectively, closing) by reconstruction, called the top-hat (respectively, inverse top-hat, or bot-hat) transform, can be used to isolate the structures that are brighter (respectively, darker) than their surroundings.

However, to determine the shapes and sizes of all objects present in the image, it is necessary to use a range of different SE sizes. This concept is called granulometry. The morphological profile (MP) of size $(2k + 1)$ can be defined as the composition of a granulometry of size k constructed with opening by reconstruction (opening profile), the original image, and an antigranulometry of size k constructed with closing by reconstruction (closing profile) using a sequence of k SEs with increasing sizes. Then, the derivative of the morphological profile (DMP) is defined as a vector where the measure of the slope of the opening-closing profile is stored for every step of an increasing SE series (see Figures 16-C and 16-D for the illustration of opening and closing profiles, respectively).



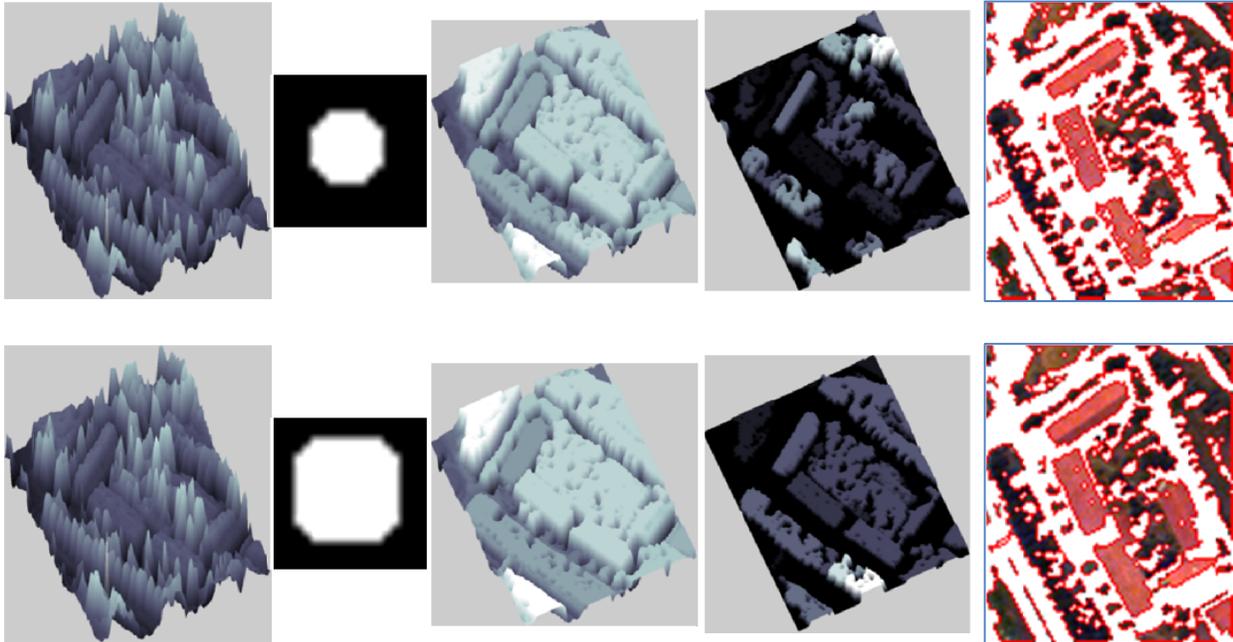
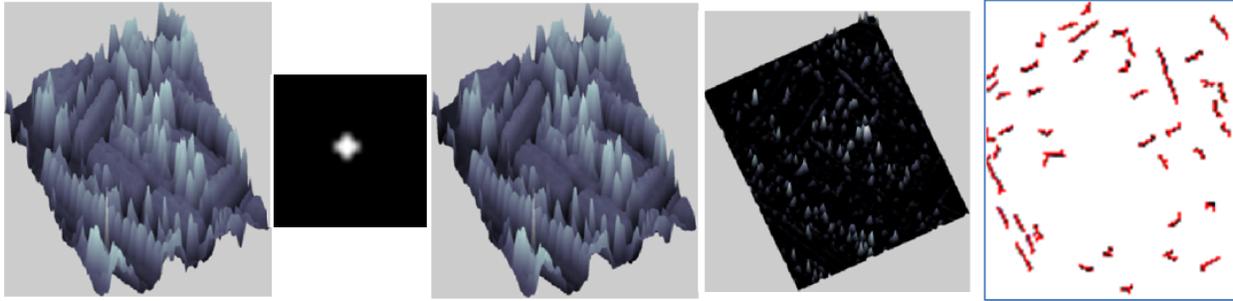


Figure 16-C. Illustration of the opening profile obtained using increasing SE sizes. Each row shows the results for an increasing SE series. The first column shows the input spectral data as a topographic relief. The second column shows the SEs used. The third column shows the result of opening by reconstruction of the topographic relief with the corresponding SEs. The fourth column shows the derivative of the opening morphological profile. The fifth column shows the boundaries of the connected components having a non-zero derivative profile for the corresponding SE for a multi-scale segmentation.



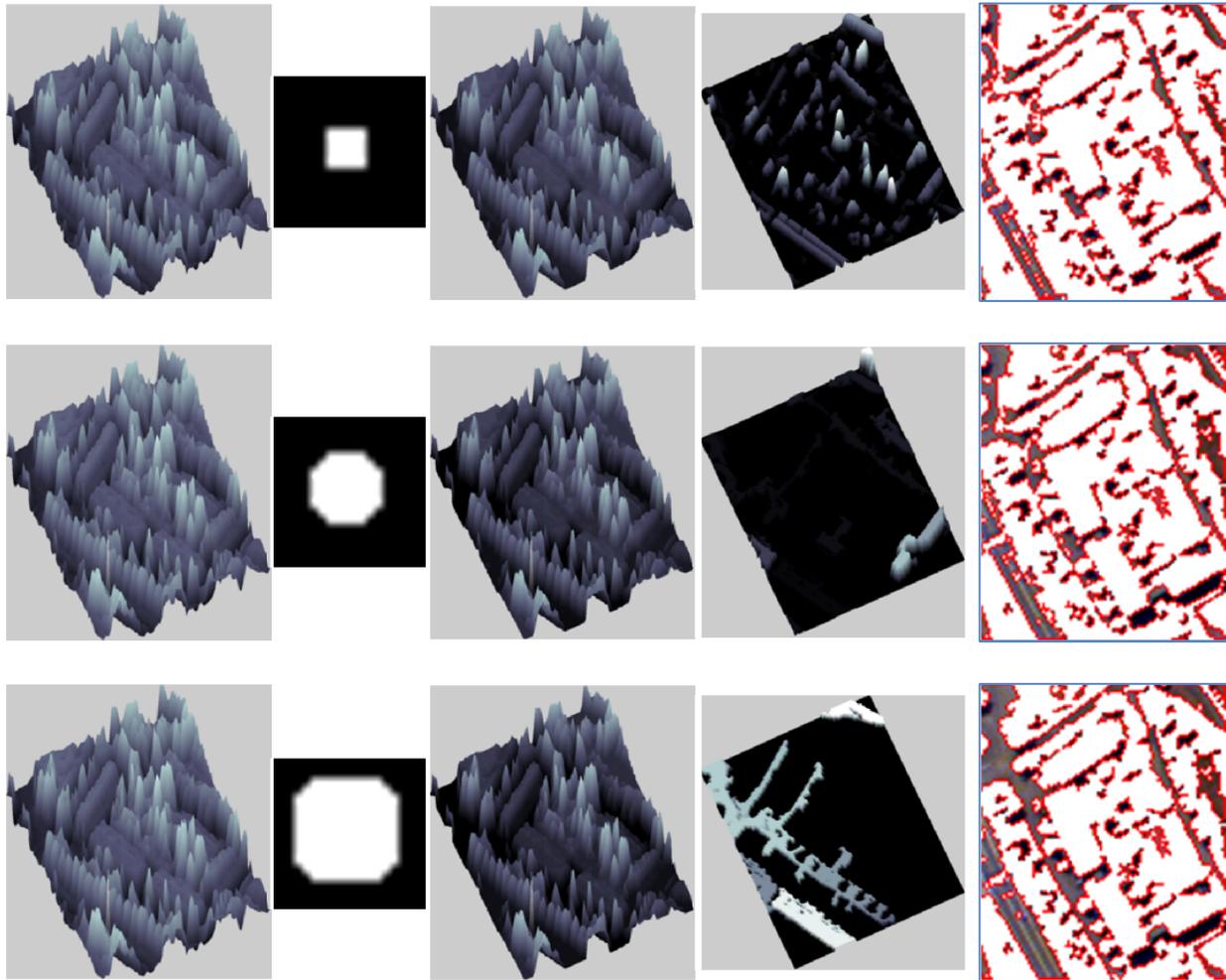


Figure 16-D. Illustration of the closing profile obtained using increasing SE sizes. Each row shows the results for an increasing SE series. The first column shows the input spectral data as a topographic relief. The second column shows the SEs used. The third column shows the result of closing by reconstruction of the topographic relief with the corresponding SEs. The fourth column shows the derivative of the closing morphological profile. The fifth column shows the boundaries of the connected components having a non-zero derivative profile for the corresponding SE for a multi-scale segmentation.

Pesaresi and Benediktsson (2001) used DMP for image segmentation. They defined the size of each pixel as the SE size at which the maximum derivative of the morphological profile is achieved. Then, they defined an image segment as a set of connected pixels showing the greatest value of the DMP for the same SE size. That is, the segment label of each pixel is assigned according to the scale corresponding to the largest derivative of its profile. This scheme works well in images where the structures are mostly flat so that all pixels in a structure have only one derivative maximum. A potential drawback of this scheme is that neighborhood information is not used at the final step of assigning segment labels to pixels. This may result in an over-segmentation consisting of small noisy segments in very high spatial resolution images with non-flat structures where the scale with the largest value of the DMP may not correspond to the true structure.

Akçay and Aksoy (2008) proposed to consider the behavior of the neighbors of a pixel while assigning the segment label for that pixel. The method assumes that pixels with a positive DMP value at a particular SE size face a change with respect to their neighborhoods at that scale. As opposed to Pesaresi and Benediktsson (2001) where only the scale corresponding to the greatest DMP is used, the main idea is that a neighboring group of pixels that have a similar change for any particular SE size is a candidate segment for the final segmentation. These groups can be found by applying connected components analysis to the DMP at each scale. For each opening and closing profile, through increasing SE sizes from 1 to m , each morphological operation reveals connected components that are contained within each other in a hierarchical manner where a pixel may be assigned to more than one connected component appearing at different SE sizes. Each component is treated as a candidate meaningful segment (see Figures 16-C and 16-D). Using these segments, a tree is constructed where each connected component is a node and

there is an edge between two nodes corresponding to two consecutive scales if one node is contained within the other. Leaf nodes represent the components that appear for SE size 1. Root nodes represent the components that exist for SE size m .

After forming a tree for each opening and closing profile, the goal is to search for the most meaningful connected components among those appearing at different scales in the segmentation hierarchy. Ideally, a meaningful segment is expected to be spectrally as homogeneous as possible. However, in the extreme case, a single pixel is the most homogeneous. Hence, a segment is also desired to be as large as possible. In general, a segment stays almost the same (both in spectral homogeneity and size) for some number of SEs, and then faces a large change at a particular scale either because it merges with its surroundings to make a new structure or because it is completely lost. Consequently, the size of interest corresponds to the scale right before this change. In other words, if the nodes on a path in the tree stay homogeneous until some node n , and then the homogeneity is lost in the next level, it can be said that n corresponds to a meaningful segment in the hierarchy. With this motivation, to check the meaningfulness of a node, Akcay and Aksoy (2008) defined a measure consisting of two factors: spectral homogeneity, which is calculated in terms of the difference of the standard deviation of the spectral features of the node and its parent, and neighborhood connectivity, which is calculated using sizes of connected components. Then, starting from the leaf nodes (level 1) up to the root node (level m), this measure is computed at each node, and a node is selected as a meaningful segment if it is highly homogeneous and large enough on its path in the hierarchy (a path corresponds to the set of nodes from a leaf to the root).

After the tree is finalized, each node is regarded as a candidate segment for the final segmentation. Given the goodness measure of each node in the hierarchy, the segments that

optimize this measure are selected by using a two-pass algorithm that satisfies the following conditions. Given N as the set of all nodes and P as the set of all paths in the tree, the algorithm selects $N^* \subseteq N$ as the final segmentation such that any node in N^* must have a measure greater than all of its descendants, any two nodes in N^* cannot be on the same path (i.e., the corresponding segments cannot overlap in the hierarchical segmentation), and every path must include a node that is in N^* (i.e., the segmentation must cover the whole image). The first pass finds the nodes having a measure greater than all of their descendants in a bottom-up traversal. The second pass selects the most meaningful nodes having the largest measure on their corresponding paths of the tree in a top-down traversal. The details of the algorithm can be found in (Akçay and Aksoy, 2008). Even though the algorithm was illustrated using a tree constructed from a DMP, it is a generic selection algorithm in the sense that it can be used with other hierarchical image partitions, such as the ones described in Section 16.3.3, and can be applied to specific applications by defining different goodness measures for desired image segments (e.g., see (Genctav, et al., 2012) for an application of this selection algorithm to a hierarchical segmentation produced by a multi-scale watershed procedure).

16.3.4. Graph-based Algorithms

Graph-based segmentation techniques gained popularity in recent years. In the graph-based framework, the image is modeled by a graph, where nodes typically represent individual pixels or regions, while edges connect spatially adjacent nodes. The weights of the edges reflect the (dis)similarity between the neighboring pixels/regions linked by the edge. The general idea is then to find subgraphs in this graph, which correspond to regions in the image scene. The early graph-theoretic approaches for image segmentation were described in (Zahn, 1971), where a minimum spanning tree was used to produce connected groups of vertices, and (Narendra, 1977),

where directed graphs have been employed to define regions in edge-detected images. In this section we will review two algorithms that have been successfully applied for remote sensing applications: optimal spanning forests and normalized cuts.

16.3.4.1. Optimal Spanning Forests

The optimal spanning forest segmentation is based on the minimum spanning tree algorithm introduced by Kruskal (1956) and Prim (1957). It was employed for segmentation of remote sensing images in (Tarabalka, et al., 2009) and (Skurikhin, 2010).

We denote an image undirected graph as $G = (V, E, W)$, where each pixel is considered as a vertex $v \in V$, each edge $e_{i,j} \in E$ connects a couple of vertices i and j corresponding to the neighboring pixels. Furthermore, a weight $w_{i,j}$ is assigned to each edge $e_{i,j}$, which indicates the degree of dissimilarity between two pixels connected by this edge. Different dissimilarity measures can be used to compute weights of edges, such as vector norms, SAM or spectral information divergence (Tarabalka, et al., 2010a).

Given a connected graph $G = (V, E)$, a spanning tree $T = (V, E_T)$ of G is a connected graph without cycles such that $E_T \subset E$. A spanning forest $F = (V, E_F)$ of G is a non-connected graph without cycles such that $E_F \subset E$. Given a graph $G = (V, E, W)$, the *minimum spanning tree* is defined as a spanning tree $T^* = (V, E_{T^*})$ of G such that the sum of the edge weights of T^* is minimal among all the possible spanning trees of G . The *minimum spanning forest* (MSF) rooted on a set of m distinct vertices $\{t_1, \dots, t_m\}$ consists in finding a spanning forest $F^* = (V, E_{F^*})$ of G , such that each distinct tree of F^* is grown from one root t_i , and the sum of the edge weights of F^* is minimal among all the spanning forests of G rooted on $\{t_1, \dots, t_m\}$.

The MSF-based segmentation typically consists of two steps:

- i. The objective of this step is to select a *marker*, or region seed, for each spatial object in the image. Such region seeds $\{t_1, \dots, t_m\}$ can be manually selected from image pixels via interactive image analysis software, however automatic marker selection is highly desirable. Markers are often defined by automatically searching flat zones (i.e. connected components of pixels of constant intensity value), zones of homogeneous texture, or image extrema (Soille, 2003). Tarabalka, et al. (2010a) proposed to perform a supervised probabilistic classification of each pixel (i.e. compute probabilities for each pixel to belong to each of the land categories of interest), and to choose the most reliably classified pixels as markers of spatial regions.
- ii. Image pixels are grouped into an MSF, where each tree is rooted on a marker. To compute an MSF, an additional root vertex r is added and is connected by the null-weight edges to the marker vertices t_i . The minimum spanning tree of the constructed graph induces an MSF in G , where each tree is grown on a marker vertex t_i ; the MSF is obtained after removing the vertex r . The two most commonly used algorithms for computing a minimum spanning tree are Prim's (1957) and Kruskal's (1956) algorithms.

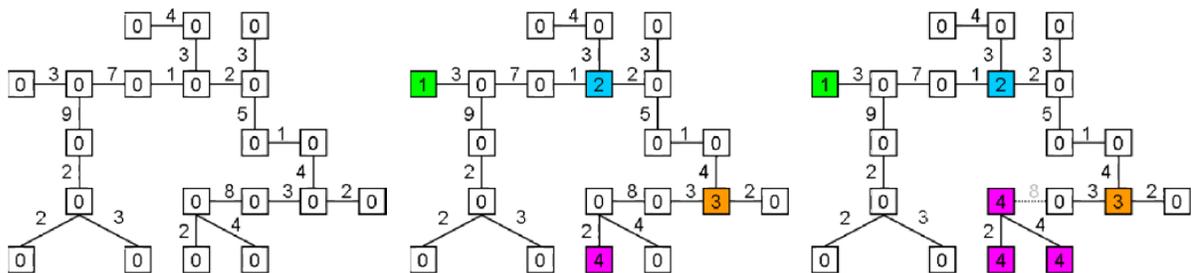
The watershed transform described in the previous section can be efficiently built by computing an MSF rooted on the image minima (Cousty, et al., 2009). For this purpose, an ultrametric flooding distance has to be used to compute weights of edges (Meyer, 2005). This distance is defined as the minimal level of flooding for which two pixels belong to the same lake.

Meyer (2005) showed that an MSF can also be efficiently computed from a minimum spanning tree of image pixels, without introducing an additional root vertex r , as depicted in Figure 16-4. This algorithm is useful if the initial markers can be modified (e.g. suppression and addition of markers during interactive segmentation). Given the minimum spanning tree T of a

graph $G = (V, E, W)$ and a set of markers $\{t_1, \dots, t_m\}$, the edges of T are first sorted in the order of their decreasing weights and are considered one after another. Suppose that e is the edge currently under consideration. The edge e belongs to a sub-tree of T . Suppressing e will cut this tree into two smaller sub-trees; if each of them contains at least one marker, then the suppression of e is validated (this is the case in Figure 16-4 (c)-(e), where an edge has been suppressed each time); if at least one of the sub-trees does not contain a marker, then the edge e is reintroduced. The process stops when each of the created sub-trees contains one and only one marker. This algorithm outputs an MSF with one tree rooted in each marker. It was applied in (Bernard, et al., 2012) for segmentation of hyperspectral remote sensing images. If the markers of the image regions cannot be reliably found, a similar algorithm can be iteratively applied, by suppressing the edge of the minimum spanning tree with the highest weight at each iteration until convergence. A threshold for the edge weight can control in this case the convergence.

16.3.4.2. Normalized Cuts

The normalized cuts segmentation method introduced by (Shi & Malik, 2000) aims at partitioning the image in the way to minimize the similarity between adjacent regions while maximizing the similarity within the regions. Figure 16-5 shows an example graph, where the pixels in group X are strongly connected with high similarities between adjacent pixels, shown



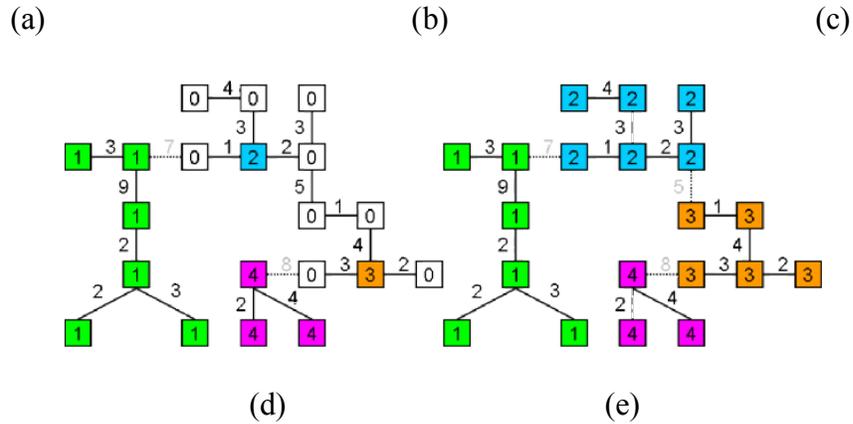


Figure 16-4. Example of construction of an MSF rooted on markers from a minimum spanning tree. (a) The initial minimum spanning tree. (b) Four markers defined by the colored nodes. (c)-(d) Illustration of the construction of the MSF from the four markers by highest weight edge suppression. (e) Final MSF, where each tree has the color of its marker.

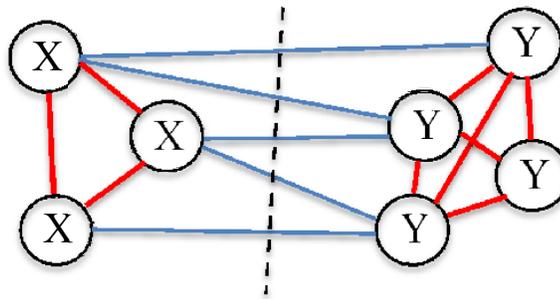


Figure 16-5. Example of a weighted graph and its normalized cut (shown as a dashed line).

as thick red lines, as are the pixels in group Y. The connections between groups X and Y, shown as blue lines, are much weaker. A normalized cut between these two groups separates them into two clusters.

The *cut* between two groups X and Y is computed as the sum of the weights of all edges being cut:

$$cut(X, Y) = \sum_{i \in X, j \in Y} w_{i,j}, \quad (16-19)$$

where the weights of the edges between two vertices i and j measure the similarity between the corresponding pixels (or regions). The optimal bipartitioning of the graph is the one that minimizes this cut value, i.e. the finds the *minimum cut* of a graph. However, because the value of the cut computed by Equation (16-19) increases with the number of edges separating two partitions, using a minimum cut as a segmentation criterion favors keeping in one group isolated pixels or small sets of isolated nodes in the graph.

To avoid such partitioning of small sets of nodes into separate groups, (Shi and Malik, 2000) proposed a new measure of disassociation between two groups. The new measure, called a *normalized cut*, computes the cut cost as a fraction of the total edge connections to all the nodes in the graph:

$$Ncut(X, Y) = \frac{cut(X,Y)}{assoc(X,V)} + \frac{cut(X,Y)}{assoc(Y,V)}, \quad (16-20)$$

where $assoc(X, V) = \sum_{i \in X, v \in V} w_{i,v}$ is the sum of all weights from nodes in X to all nodes in the graph, and $assoc(Y, V)$ is similarly defined. The normalized cut defined by Equation (16-20) better reflects the fitness of a particular segmentation, because it seeks to cut a collection of edges that are weak relatively to all of the edges both inside and emanating from each of the regions. Shi and Malik (2000) proposed an efficient algorithm based on a generalized eigenvalue problem to optimize the normalized cut criterion.

Numerous works employed a normalized cut algorithm for segmentation of remote sensing images, such as (Grote, et al., 2007) and (Sung and He, 2009). Normalized cut were also applied in combination with other techniques. For instance, Jing, et al. (2010) first used watershed algorithm to find initial segments, and then the normalized cuts technique grouped these segments into the final regions.

16.3.5. MRF-based Algorithms

Markov random fields (MRFs) are probabilistic graphical models that conceptually generalize the notion of Markov chain (Wang, et al., 2013), (Moser, et al., 2013). They provide a flexible tool to include spatial context into image-analysis schemes in terms of minimization of suitable energy functions. While earlier algorithms for optimizing MRF energy, such as iterated conditional modes (ICM) and simulated annealing (Solberg, et al., 1996), (Tarabalka, et al., 2010b) were time consuming, more advanced methods, such as graph cuts (Boykov, et al., 2001), (Li, et al., 2012) provided powerful alternatives from both theoretical and computational viewpoints, resulting in a growing use of the MRF-based segmentation techniques.

For land categorization applications, MRFs are usually applied in the framework of image classification, where the output is a landcover map, with every region assigned to one of the thematic classes. The commonly used MRF energy function in this case is computed as a linear combination of two terms:

$$E(L) = E_{data} + E_{smooth}. \quad (16-21)$$

The first term $E_{data} = \sum_{i=1}^n V_i(L_i)$ is related to pixelwise information and it measures for each pixel the disagreement between a prior probabilistic model and the observed data. Thus, individual potentials $V_i(L_i)$ measure a penalty for a pixel i ($i = 1, 2, \dots, n$) to have a label L_i . This term is often formulated in terms of the probability density function of feature vectors conditioned to the related class label:

$$V_i(L_i) = -\ln(p(x_i|L_i)). \quad (16-22)$$

This probability density can be estimated based on the available training samples for each landcover class. A discussion of the main approaches for this estimation problem can be found in (Duda, et al., 2001)

The second contribution $E_{smooth} = \sum_{i \sim j} W_{i,j}(L_i, L_j)$ expresses interaction between neighboring pixels, thus exploiting image spatial context. $i \sim j$ denotes a pair of spatially adjacent pixels, and $W_{i,j}(L_i, L_j)$ is an interaction term for these pixels. Most works in remote sensing image classification use a Potts model (Tarabalka, et al., 2010b), (Li, et al., 2012) to compute this spatial term, which favors spatially adjacent pixels to belong to the same class (or spatial region):

$$W_{i,j}(L_i, L_j) = \beta(1 - \delta(L_i, L_j)), \quad (16-23)$$

where $\delta(\cdot)$ is the Kronecker function ($\delta(a, b) = 1$ for $a = b$ and $\delta(a, b) = 0$ otherwise) and β is a positive constant parameter that controls the importance of spatial smoothing. This model tends to deteriorate classification results at the edges between land-cover classes and near small-scale details. In order to preserve the border in the output thematic map, "edge" functions have been proposed and integrated in the spatial energy term, such as (Tarabalka, et al., 2010b):

$$W_{i,j}(L_i, L_j) = \beta(1 - \delta(L_i, L_j)) \frac{t}{t + |\rho_{ij}|}, \quad (16-24)$$

where ρ_{ij} is the gradient value of the pixel i in the direction of j , and t is a parameter controlling the fuzzy edge threshold.

To optimize the MRF energy, different methods were proposed and applied for remote sensing applications, as described in the following.

16.3.5.1. Simulated Annealing and Iterated Conditional Modes

The main idea of simulated annealing described in (Kirkpatrick, et al., 1983) is to iteratively propose a change from the current configuration of pixel labels (region or class labels) L to the new randomly generated configuration of labels, and to probabilistically decide if this change is accepted or not. This procedure yields configurations of labels with the lower MRF energy.

Stewart, et al. (2000) applied MRF-based method with simulated annealing optimization for segmentation and classification of synthetic aperture radar (SAR) data. They used Gamma distribution to model the data energy term and a Potts model to compute the spatial term. Tarabalka, et al. (2010b) applied simulated annealing technique for classification of hyperspectral remote sensing images. They performed probabilistic support vector machines (SVM) classification to derive the data term, and employed the spatial energy contribution with the fuzzy edge function described previously. At each iteration, a new class label L_i^{new} was randomly selected for a randomly selected pixel x_i , and a new local energy $E^{new}(x_i)$ was computed. If the variation of the energy $\Delta E = E^{new}(x_i) - E(x_i) < 0$, the new class label is accepted. Otherwise, the new class assignment is accepted with the probability $p = \exp(-\Delta E/T)$, where T is a global parameter called temperature. The optimization begins at a high temperature, which is gradually lowered as the iteration procedure proceeds. This helps to avoid converging to local minima.

The iterated conditional modes (ICM) optimization algorithm proposed in (Besag, 1986) is conceptually simpler and computationally less expensive. It iteratively considers every pixel x_i , and assigns a (region or class) label to this pixel, which minimizes the local energy centered at x_i . Solberg, et al. (1996) applied the ICM optimization for multisource classification of optical, SAR and geographic information systems images. Farag, et al. (2005) employed the ICM method for hyperspectral image classification. The density functions in this case were estimated by mean field-based SVM regression. The drawback of the ICM algorithm is that it is suboptimal and converges only to a local minimum of the energy function.

16.3.5.2. Graph-cuts

Boykov and Jolly (2001) were the first to apply the graph-cut optimization technique proposed by (Greig, et al., 1989) to binary image segmentation. This algorithm is computationally efficient and it gives a globally-optimal solution to the binary segmentation problem. In the approach of (Boykov and Jolly, 2001), all pixels are connected by t -links to the additional two nodes, called source and sink, respectively, as depicted in Figure 16-6. The weights of these edges are computed as the potentials $V_i(L_i)$ (see Eq. 16-12), so that pixels that are more compatible with the foreground or background region get stronger connections to the respective source or sink. The e -links between adjacent pixels are assigned weights $W_{i,j}(L_i, L_j)$. The resulting MRF energy (see Eq. 16-21) is optimized by solving the minimum-cut/maximum-flow problem, yielding a binary segmentation map.

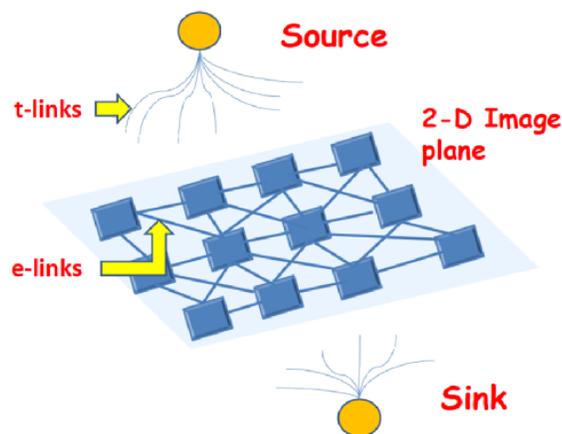


Figure 16-6. Mapping of an image to the graph for the graph-cut optimization.

Boykov, et al. (2001) proposed two efficient approximation algorithms based on graph-cuts (α -expansion and α - β -swap) to solve multi-label classification problems. In remote sensing, Denis, et al. (2009) applied a graph-cut-based algorithm for regularization of SAR images. Li, et al. (2012) used an α -expansion optimization of the MRF energy for segmentation of hyperspectral images.

16.4. Image Segmentation Quality Evaluation

We noted in the introduction to this chapter that there is no general theory of image segmentation, and most image segmentation approaches are *ad hoc* in nature. This is also the case with image segmentation quality evaluation. In practice, image segmentation quality is best evaluated by how well the goals of the analysis utilizing the segmentation are served.

Several general approaches to quantitative segmentation quality evaluation are listed and described in the series of papers published by the research group based at the Leibniz Institute for Ecological and Regional Development (IOER) mentioned in the introduction to this chapter, in particular (Neubert, et al., 2006). Among the evaluation approaches discussed and employed are:

- i. Fragmentation and Area-Fit-Index (Strasters and Gerbrands, 1991), (Lucieer, 2004): These measures address over- and under-segmentation by analyzing the number of segmented and reference regions.
- ii. Geometric features Circularity and Geometric features Shape Index (Yang, et al., 1995), (Neubert and Meinel, 2003): These measures address the shape conformity between segmentation and reference regions.
- iii. Empirical Evaluation Function (Borsotti, et al., 1998): This measure addresses how uniform a feature is within segmented regions.
- iv. Entropy-based evaluation function and a weighted disorder function (Zhang and Gerbrands, 1994): This measure addresses the uniformity within segmented regions using entropy as a criterion of disorder.
- v. Fitness function (Everingham, et al., 2002): This measure addresses multiple criteria and parameterizations of algorithms by a probabilistic Fitness/Cost Analysis.

Neubert, et al. (2006) also note that the vast majority of quantitative image segmentation approaches are empirical discrepancy methods that analyze the number of misclassified pixels in relation to a reference classification. In the remainder of this section we describe and demonstrate such an approach to image segmentation quality evaluation. The first step in this evaluation approach is to perform a pixelwise classification of an image data set. In our example we create our pixelwise classification using the support vector machine (SVM) classifier. Then a region classification is obtained by assigning each spatially connected region from the segmentation result to the most frequently occurring class within the region. The SVM classifier and this “plurality vote” (PV) region-based classification approach are described in more detail in (Tilton, et al., 2012).

Our test data set is the University of Pavia data set was recorded by the Reflective Optics System Imaging Spectrometer (ROSIS) over the University of Pavia, Pavia, Italy. The image is 610×340 pixels in size, with a spatial resolution of 1.3 m. The ROSIS sensor has 115 spectral channels, with a spectral range of 0.43-0.86 μm . The 12 noisiest channels were removed, and the remaining 103 spectral bands were used in this experiment. See (Tarabalka, et al., 2010) for more details on this data set. The reference data contain nine ground cover classes: asphalt, meadows, gravel, trees, metal sheets, bare soil, bitumen, bricks and shadows. A three-band false color image of this data set and the ground reference data are shown in Fig 16-5.

The classification accuracy results are listed in Table 16-2 for the pixelwise SVM classification and the PV region-based classification approach for several region growing approaches. Fig. 16-5 also shows the corresponding classification maps. The classification maps for the PV region-based classification approach appear smoother than the pixelwise SVM classification, with the D8 and SEGEN based classification appearing the smoothest. All of the

PV region-based classification accuracies as substantially higher than the accuracy of the SVM pixelwise classification, with the D8, SEGEN and HSeg PV region-based classifications notably more accurate than the PV region-based classifications based on Muerle Allen or HSWO segmentation. The HSeg based classification has a marginally higher classification than all the other approaches.

The PV region-based classifications are evaluated in terms of Overall Accuracy (OA), Average Accuracy (AA) and the kappa coefficient (κ). OA is the percentage of correctly classified pixels, AA is the mean of the class-specific accuracies, and κ is the percentage of agreement (correctly classified pixels) corrected by the number of agreements that would be expected purely by chance (e.g., see Richards and Jia, 1999).

Table 16-2. Comparison of classification accuracies on the University of Pavia hyperspectral data set for per-pixel SVM and with the plurality vote (PV) region-based classification method for Muerle Allen, HSWO, Definiens 8.0 (D8), SEGEN, and HSeg. HSeg was performed with small region merge acceleration. Percentage classification accuracies in terms of OA, AA, and Kappa Coefficient (κ).

	SVM	Muerle Allen	HSWO +PV	D8+PV	SEGEN +PV	HSeg+PV $S_w = 0.3$
OA	89.03	95.35	95.38	97.54	98.09	98.35
AA	89.56	95.26	95.50	97.26	97.95	98.15
κ	85.46	93.78	93.83	96.71	97.45	97.79

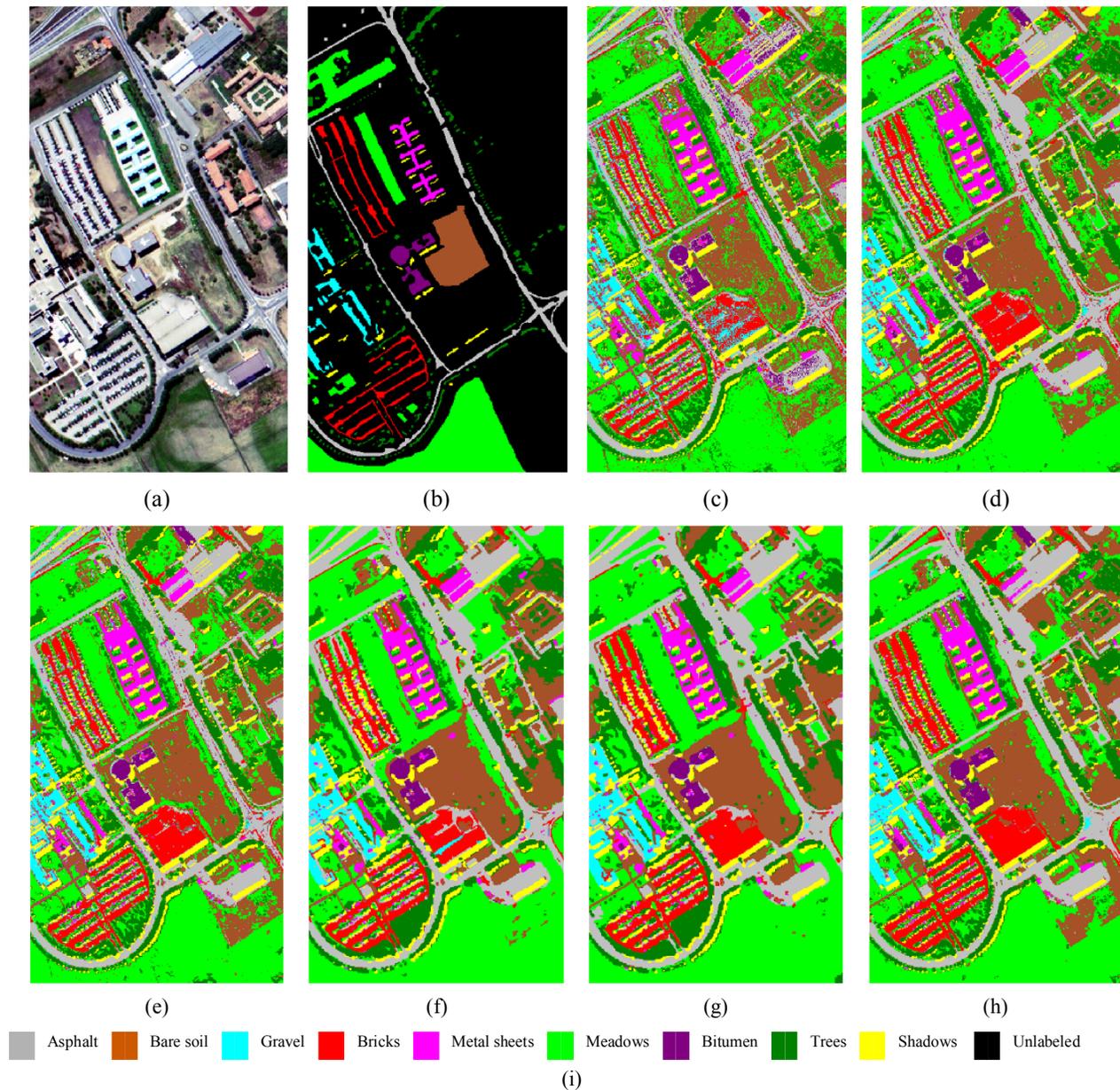


Fig. 16-5. (a) Three-band false color image of the University of Pavia hyperspectral data set (RGB = bands 56, 33 and 13). (b) Reference data. (c) SVM classification. (d) PV classification with Muerle Allen segmentation. (e) PV classification with HSWO segmentation. (f) PV classification with D8. (g) PV classification with SEGEN segmentation. (h) PV classification with HSeg segmentation ($S_w = 0.3$). (i) Color key. The training data for the SVM classifier and segmentation optimization are two separate randomly selected subsets of the reference data.

Table 16-3 and Fig. 16-6 show classification results obtained by using the graph-based segmentation approaches. The MSF segmentation with plurality vote was performed as described in (Tarabalka, et al., 2010a), by using the $L1$ norm between spectral vectors to compute weights of edges between adjacent pixels. For the graph-cut method, probabilistic SVM classification was applied to derive the data energy term, and a Potts model was used to express the spatial term. We fixed the parameter $\beta = 1.5$ as suggested in (Tarabalka, et al., 2010b). The MRF-based method using the α -expansion optimization yields the highest overall classification accuracy.

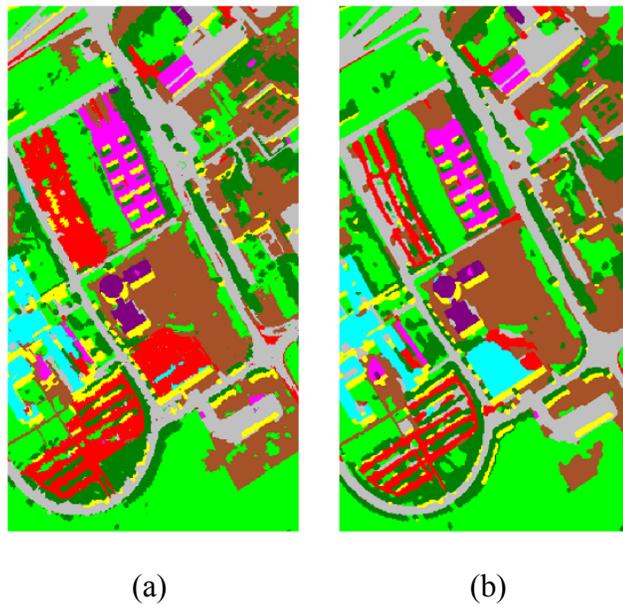


Fig. 16-6. (a) Minimum spanning forest (MSF) with plurality vote (PV) classification map. (b) Classification map of the α -expansion graph-cut algorithm.

Table 16-3. Comparison of classification accuracies on the University of Pavia hyperspectral data set for the minimum spanning forest (MSF) with plurality vote (PV) method and MRF-based method using the α -expansion graph-cut algorithm. Percentage classification accuracies in terms of OA, AA, and Kappa Coefficient (κ).

	MSF+PV	Graph-cut
OA	96.99	98.49
AA	97.01	97.73
κ	95.99	98.47

16.5. Concluding Remarks

Land categorization applications are quite varied. They range from land cover monitoring (e.g., croplands, forests, wetlands, and urbanization), snow and ice mapping, geology and mineral exploration, and even wildfire and agricultural burn monitoring. Each land categorization will have its particular image analysis needs that will be best served by a particular class of image segmentation algorithm. Generally, region-growing approaches may be best suited for finding homogenous regions, while if locating region borders are more important, watershed approaches may be more effective. Region-growing and morphological approaches that produce segmentation hierarchies would be most appropriate if analysis of the image data at different scales is desired.

We have described and discussed a wide range of image segmentation approaches developed over the past century, focusing on those approaches most applicable to the analysis of remotely sensed imagery data for land categorization. The wide range of approaches included in our discussion is a clear indication that there still is no general theory of segmentation and that most of the successful image segmentation approaches are rather *ad hoc* in nature. But the wide range of approaches provide a rich menu to choose from for image analysis practitioners in tailoring their image analysis approach to their particular application.

References

- Adams, R., and L. Bischof. 1994. Seeded region growing. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(6):641-646.
- Akcay, H. G. and S. Aksoy. 2008. Automatic detection of geospatial objects using multiple hierarchical segmentations. *IEEE Trans. Geosci. Remote Sens.*, 46(7):2097-2111.
- Akcay, H. G. and S. Aksoy. 2011. Detection of compound structures using hierarchical clustering of statistical and structural features. *Proc. of IGARSS*, Vancouver, Canada.
- Akcay, H. G., S. Aksoy, and P. Soille. 2010. Hierarchical segmentation of complex structures. *Proc. of ICPR*, Istanbul, Turkey.
- Aksoy, S. 2006. Spatial techniques for image classification. In: C. H. Chen, ed. *Signal and Image Processing for Remote Sensing* Boca Raton, FL: CRC Press, pp. 491-513.
- Aksoy, S., and H. G. Akcay. 2005. Multi-resolution segmentation and shape analysis for remote sensing image classification. *Proc. of 2nd Int'l Conf. Recent Advances in Space Technologies*, Istanbul, Turkey.
- Aksoy, S., H. G. Akcay, and T. Wassenaar. 2010. Automatic mapping of linear woody vegetation features in agricultural landscapes using very-high resolution imagery. *IEEE Trans. Geosci. Remote Sens.*, 48(1):511-522.
- Aksoy, S., K. Koperski, C. Tusk, G. Marchisio, and J. C. Tilton. 2005. Learning Bayesian classifiers for scene classification with a visual grammar. *IEEE Trans. Geosci. Remote Sens.*, 43(3):581-589.
- Aksoy, S., I. Z. Yalniz, and K. Tasdemir, 2012. Automatic detection and segmentation of orchards using very high-resolution imagery. *IEEE Trans. Geosci. Remote Sens.*, 50(8):3117-3131.

- Aptoula, E., and S. Lefevre. 2007. A comparative study on multivariate mathematical morphology. *Pattern Recognition*, 40(11):2914-2929.
- Baatz, M., and A. Schape. 2000. Multiresolution segmentation: an optimizing approach for high quality multi-scale segmentation. In: J. Strobl and T. Blaschke, eds. *Angewandte Geographich Informationsverarbeitung, XII*. Heidelberg, Germany: Wichmann, pp. 12-23.
- Beaulieu, J.-M., and M. Goldberg. 1989. Hierarchy in picture segmentation: A stepwise optimal approach. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11:150-163.
- Benz, U., P. Hofmann, G. Wilhauck, I. Lingenfelder, and M. Heynen. 2004. Multi-resolution, objected-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS J. Photogramm. Remote Sens.*, 58(3/4):239-258.
- Bernard, K., Y. Tarabalka, J. Angulo, J. Chanussot, and J. A. Benediktsson. 2012. Spectral-spatial classification of hyperspectral data based on a stochastic minimum spanning forest approach. *IEEE Trans. Image Proc.*, 21(4):2008-2021.
- Besag, J. 1986. On the statistical analysis of dirty pictures. *J. R. Statist. Soc.*, 48(3):259-302.
- Borsotti, M., P. Campadelli, and R. Schettini. 1998. Quantitative evaluation of color image segmentation results. *Pattern Recognition Letters*, 19(8):741-747.
- Boykov, Y., M.-P. Jolly. 2001. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. *Proc. of ICCV*, Vancouver, Canada, 1:105-112.
- Boykov, Y., O. Veksler, and R. Zabih. 2001. Fast approximate energy minimization via graph cuts. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(11):1222-1239.
- Bruzzone, L., and D. F. Prieto, 2005. Automatic analysis of the difference image for unsupervised change detection. *IEEE Trans. Geosci. Remote Sens.*, 38(3):1171-1182.

- Castilla, G., Hay, G. J. and Ruiz, J. R., 2008. Size-constrained region merging (SCRM): an automated delineation tool for assisted photointerpretation. *PE&RS*, 74(4):409-419.
- Comaniciu, D., and P. Meer. 2002. Mean shift: a robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(5):603-619.
- Cook, R., I. McConnell, D. Stewart and C. J. Oliver. 1996. Segmentation and simulated annealing. In: G. Franceschetti, F. S. Rubertone, C. J. Oliver, and S. Tajbakhsh, eds. *Microwave Sensing and Synthetic Aperture Radar, Proc. SPIE* 2958:30-35.
- Couprie, C., L. Grady, L. Najman, and H. Talbot. 2011. Power watershed: A unifying graph-based optimization framework. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(7):1384-1399.
- Cousty, J., G. Bertrand, L. Najman, and M. Couprie. 2009. Watershed cuts: Minimum spanning forests and the drop of water principle. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(8):1362-1374.
- Cross, A. M., D. C. Mason and S. J. Dury. 1988. Segmentation of remotely sensed image by a split-and-merge process. *Int. J. Remote Sens.*, 9:1329-1345.
- Denis, L., F. Tupin, J. Darbon, and M. Sigelle, 2009. SAR image regularization with fast approximate discrete minimization. *IEEE Trans. Image Proc.*, 18(7):1588-1600.
- Dogrusoz, E., and S. Aksoy. 2007. Modeling urban structures using graph-based spatial patterns. *Proc. of IGARSS*, Barcelona, Spain, pp. 4826-4829.
- Dongping, M., C. Tianyu, C. Hongyue, L. Longxiang, Q. Cheng, and D. Jinyang. 2012. Semivariogram-based spatial bandwidth selection for remote sensing image segmentation with mean-shift algorithm. *IEEE Geoscience and Remote Sensing Letters*, 9(5): 813-817.
- Duda, R. O., P. E. Hart, and D. G. Stork. 2001. *Pattern classification, 2nd ed.*. New York: Wiley.

- Epifanio, I., P. Soille. 2007. Morphological texture features for unsupervised and supervised segmentations of natural landscapes. *IEEE Trans. Geosci. Remote Sens.*, 45(4):1074-1083.
- Everingham, M., H. Muller, and B. Thomas. 2002. Evaluating image segmentation algorithms using monotonic hulls in fitness/cost space. In: T. Cootes and C. Taylor. (Eds.): *Proc. 12th British Machine Vision Conference*, pp. 363-372.
- Farag, A. A., R. M. Mohamed, and A. El-Baz. 2005. A unified framework for MAP estimation in remote sensing image segmentation. *IEEE Trans. Geosci. Remote Sens.*, 43(7):1617-1634.
- Fauvel, M., Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton. 2013. Advances in spectral-spatial classification of hyperspectral images. *Proceedings of the IEEE*, 101(3):652-675.
- Fu, K. S., and J. K. Mui. 1981. A survey on image segmentation. *Pattern Recognition*, 13:3-16.
- Gaetano, R., G. Scarpa, and G. Poggi. 2009. Hierarchical texture-based segmentation of multiresolution remote-sensing images. *IEEE Trans. Geosci. Remote Sens.*, 47(7):2129-2141.
- Genctay, A., S. Aksoy, and S. Onder S. 2012. Unsupervised Segmentation and Classification of Cervical Cell Images. *Pattern Recognition*, 45(12):4151-4168.
- Ghamisi, P., M. S. Couceiro, F. M. L. Martins, and J. A. Benediktsson. 2014. Multilevel image segmentation based on fractional-order Darwinian particle swarm optimization. *IEEE Trans. Geosci. Remote Sens.*, 52(5):2382-2394.
- Gofman, E. 2006. Developing an efficient region growing engine for image segmentation. *Proc. of ICPR*, Hong Kong, China, pp. 2413-2416.

- Greig, D., B. Porteous, and A. Seheult. 1989. Exact maximum a posteriori estimation for binary images. *Journal of the Royal Statistical Society, Series B*, 51(2):271-279.
- Grote, A., M. Betenuth, M. Gerke, and C. Heipke, 2007. Segmentation Based on Normalized Cuts for the Detection of Suburban Roads in Aerial Imagery. *Urban Remote Sensing Joint Event*, Paris, France, pp. 1-5.
- Gueguen, L., P. Soille, and M. Pesaresi. 2012. A new built-up presence index based on density of corners. *Proc. of IGARSS*, Munich, Germany.
- Haralick, R. M., and L. G. Shapiro. 1985. Survey: Image segmentation techniques. *Computer Vision, Graphics, and Image Processing*, 29(1):100-132.
- Haris, K., S. N. Efstratiadis, N. Maglaveras, and A. Katsaggelos. 1998. Hybrid image segmentation using watersheds and fast region merging. *IEEE Trans. Image Proc.*, 7(12): 1684-1699.
- Horowitz, S. L., and T. Pavlidis. 1974. Picture segmentation by a directed split-and-merge procedure. *Proc. of the 2nd Int'l Joint Conf. Pattern Recognition*, pp. 424-433.
- Jing, W., J. Hua, and W. Yubin. 2010. Normalized cut as basic tool for remote sensing image. *Proc. of Int'l Conf. on ICISS*, Gandhinagar, India, pp. 247-249.
- Kirkpatrick, S., C. D. Gelatt Jr, and M. P. Vecchi. 1983. Optimization by Simulated Annealing. *Science*, 220(4598):671-680.
- Kruse, F. A., A. B. Lefkoff, J. W. Boardman, K. B. Heidebrecht, A. T. Shapiro, P. J. Barloon, and A. F. H. Goetz. 1993. The spectral image processing system (SIPS) - Interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ.*, 44(2/3):145-163.

- Kruskal, J. 1956. On the shortest spanning tree of a graph and the traveling salesman problem. *Proc. of Am. Math. Soc.*, 7:48-50.
- Kurita, T. 1994. An efficient agglomerative clustering algorithm for region growing. Proc. of MVA, IAPR Workshop on Mach. Vis. Appl., Kawasaki, Japan, pp. 210-213.
- Kurtz, C., N. Passat, P. Gancarski, and A. Puissant. 2012. Extraction of complex patterns from multiresolution remote sensing images: A hierarchical top-down methodology. *Pattern Recognition*, 45(2):685-706.
- Landgrebe, D. A. 2003. *Signal Theory Methods in Multispectral Remote Sensing*, Hoboken, NJ: Wiley.
- Le Moigne, J. and J. C. Tilton. 1995. Refining image segmentation by integration of edge and region data. *IEEE Trans. Geosci. Remote Sens.*, 33(3):605-615.
- Li, J., J. M. Bioucas-Dias, and A. Plaza. 2012. Spectral-Spatial Hyperspectral Image Segmentation Using Subspace Multinomial Logistic Regression and Markov Random Fields. *IEEE Trans. Geosci. Remote Sens.*, 50(3):809-823.
- Li, P., and B. Xiao. 2007. Multispectral image segmentation by a multichannel watershed-based approach. *Int'l Jour. Remote Sens.*, 28(19):4429-4452.
- Lucieer, A., 2004. Uncertainties in Segmentation and their Visualisation, PhD Thesis, International Institute for Geo-Information Science and Earth Observation (ITC) and University of Utrecht, The Netherlands.
- Marpu, P. R., M. Neubert, H. Herold, and I. Niemeyer. 2010. Enhanced evaluation of image segmentation results. *Jour. Spatial Science*, 55(1):55-68.

- Meinel, G. and M. Neubert, M., 2004. A comparison of segmentation programs for high resolution remote sensing data. *Proc. of Commission IV, XXth ISPRS Congr.*, Istanbul, Turkey, pp. 1097-1102.
- Meyer, F. 2001. An overview of morphological segmentation. *Int'l Jour. of Pattern Recognition and Artificial Intelligence*, 15(7):1089-1118.
- Meyer, F., 2005. Grey-weighted, ultrametric and lexicographic distances. In: C. Ronse, L. Najman, and Etienne Decenciere, eds, *Mathematical Morphology: 40 Years On*. Netherlands:Springer, pp. 289-298.
- Meyer, F., and S. Beucher. 1990. Morphological segmentation. *Journal of Visual Communication and Image Representation*, 1(1):21-46.
- Moser, G., S. B. Serpico, and J. A. Benediktsson. 2013. Landcover mapping by Markov modeling of spatial-contextual information in very-high-resolution remote sensing images. *Proc. of IEEE*, 101(3):631-651.
- Muerle, J. L., and D. C. Allen, 1968. Experimental evaluation of techniques for automatic segmentation of objects in a complex scene. In: G. C. Cheng, ed. *Pictorial Pattern Recognition*. Washington, DC: Thompson, pp. 3-13.
- Najman, L., and M. Schmitt. 1996. Geodesic saliency of watershed contours and hierarchical segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 18(12): 1163-1173.
- Narendra, P. M., and M. Goldberg. 1977. A graph-theoretic approach to image segmentation. *Proc. of IEEE Comp. Soc. Conf. Pattern Recognition Image Proc.*, pp. 248-256.
- Neubert, M., H. Herold, and G. Meinel. 2006. Evaluation of remote sensing image segmentation quality - Further results and concepts. *Proc. 1st Int. Conf. OBIA*, Salzburg, Austria.

- Neubert, M., H. Herold, and G. Meinel. 2008. Assessing image segmentation quality - Concepts, methods and application. In: T. Blaschke, S. Lang and G. J. Hay, eds. *Object-Base Image Analysis: Spatial Concepts for Knowledge-Driven Remote Sensing Applications*. Berlin, Germany: Springer-Verlag.
- Neubert, M., and G. Meinel. 2003. Evaluation of segmentation programs for high resolution remote sensing applications. In: Proc. Joint ISPRS/EARSel. Workshop "High Resolution Mapping from Space 2003". Hannover, Germany.
- Noyel, G., J. Angulo, and D. Jeulin. 2007. Morphological segmentation of hyperspectral images. *Image Analysis and Stereology*, 26(3):101-109.
- Otsu, N., 1979. A threshold selection method from gray-level histograms. *IEEE Trans. Systems, Man, and Cybernetics*, SMC-9(1):62-66.
- Perret, B., S. Lefevre, C. Collet, and E. Slezak. 2012. Hyperconnections and hierarchical representations for grayscale and multiband image processing. *IEEE Trans. Image Proc.*, 30(7):14-27.
- Pesaresi, M., J. A. Benediktsson. 2001. A new approach for the morphological segmentation of high-resolution satellite imagery. *IEEE Trans. Geosci. Remote Sens.*, 39(2):309-320.
- Pesaresi, M., A. Gerhardinger, and F. Kayitakire. 2008. A robust built-up area presence index by anisotropic rotation-invariant textural measure. *IEEE JSTARS*, 1(3):180-192.
- Plaza, A. J., and J. C. Tilton. 2005. Automated selection of results in hierarchical segmentations of remotely sensed hyperspectral images. *Proc. of IGARSS*, Seoul, Korea, 7:4946-4949.
- Prim, R., 1957. Shortest connection networks and some generalizations. *Bell Systems Technical J.*, 36:1389-1401.

- Richards, J. A. and X. Jia. 1999. *Remote Sensing Digital Image Analysis: An Introduction*, New York: Springer-Verlag.
- Robinson, D. J., N. J. Redding, and D. J. Crisp. 2002. Implementation of a fast algorithm for segmenting SAR imagery. In: *Scientific and Technical Report*. Defence Science and Technology Organization, Edinburgh, South Australia, Australia.
- Rosin, P. L., and J. Hervas. 2005. Remote sensing image thresholding methods for determining landslide activity. *Int'l Jour. Remote Sensing*, 26(6):1075-1092.
- Sezgin, M., and B. Sankur. 2004. Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1):146-168.
- Shankar, B. U. (2007). Novel classification and segmentation techniques with application to remotely sensed images. In: V. W. Marek, E. Orłowska, R. Slowinski, and W. Ziarko, eds., *Transactions on Rough Sets VII*, New York:Springer-Verlag, pp. 295-380.
- Shi, J., and J. Malik. 2000. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(8): 888-905.
- Sirmacek, B., and C. Unsalan. 2010. Urban area detection using local feature points and spatial voting. *IEEE Geoscience and Remote Sensing Letters*, 7(1): 146-150.
- Skurikhin, A. N., 2010. Patch-based image segmentation of satellite imagery using minimum spanning tree construction. *Proc. GEOBIA*, Ghent, Belgium.
- Soille, P. 2003. *Morphological Image Analysis*. New York: Springer-Verlag.
- Soille, P. 2008. Constrained connectivity for hierarchical image partitioning and simplification. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(7):1132-1145.

- Soille, P., and L. Najman. 2010. On morphological hierarchical representations for image processing and spatial data clustering. *Int'l Workshop on Applications of Discrete Geometry and Mathematical Morphology*, Istanbul, Turkey, pp. 43-67.
- Soille, P. and Pesaresi, M., 2002. Advances in mathematical morphology applied to geoscience and remote sensing. *IEEE Trans. Geosci. Remote Sens.*, 40(9): 2042-2055.
- Solberg, A. H. S., T. Taxt, and A. K. Jain, 1996. A Markov random field model for classification of multisource satellite imagery. *IEEE Trans. Geosci. Remote Sens.*, 34(1):100-113.
- Sramek, M., T. Wrška, 1997. Watershed based image segmentation - an effective tool for detecting landscape structure. *Digital Image Processing and Computer Graphics (DIP'97)*, Proc. SPIE 3346, pp. 227-235.
- Stewart, D., B. Blacknell, A. Blake, R. Cook, and C. Oliver , 2000. Optimal approach to SAR segmentation and classification. *IEE Proceedings on Radar, Sonar and Navigation*, 147(3):pp. 134-142.
- Strasters, K., and J. Gerbrands. 1991. Three-dimensional segmentation using a split, merge and group approach. *Pattern Recognition Letters*, 12:307-325.
- Sung, F., J. He. 2009. The remote-sensing image segmentation using textons in the Normalized Cuts framework. *Mechatronics and Automation: International Conference on ICMA*, Changchun, China, pp. 1877-1881.
- Tarabalka, Y., J. A. Benediktsson, J. Chanussot, J. C. Tilton. 2010. Multiple spectral-spatial classification approach for hyperspectral data. *IEEE Trans. Geosci. Remote Sens.*, 48(11):4122-4132.

- Tarabalka, Y., J. Chanussot, J. A. Benediktsson. 2009. Classification of hyperspectral images using automatic marker selection and minimum spanning forest. *Proc. of IEEE WHISPERS*, Grenoble, France, pp. 1-4.
- Tarabalka, Y., J. Chanussot, J. A. Benediktsson., 2010a. Segmentation and Classification of Hyperspectral Images Using Minimum Spanning Forest Grown From Automatically Selected Markers. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 40(5):1267-1279.
- Tarabalka, Y., M. Fauvel, J. Chanussot, J. A. Benediktsson. 2010b. SVM- and MRF-Based Method for Accurate Classification of Hyperspectral Images. *IEEE Geosc. Remote Sensing Letters*, 7(4):736-740.
- Tarabalka, Y., J. C. Tilton, J. A. Benediktsson, and J. Chanussot. 2012. A Marker-Based Approach for the Automated Selection of a Single Segmentation From a Hierarchical Set of Image Segmentations. *IEEE JSTARS*, 5(1):262-272.
- Tilton, J. C., 1998. Image segmentation by region growing and spectral clustering with a natural convergence criterion. *Proc. of IGARSS*, Seattle, WA, USA.
- Tilton, J. C., 2003. Analysis of hierarchically related image segmentations. *Proc. of IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, Greenbelt, MD, pp. 60-69.
- Tilton, J. C., 2007. Parallel implementation of the recursive approximation of an unsupervised hierarchical segmentation algorithm. In: A. J. Plaza and C. Chang, eds. *High Performance Computing in Remote Sensing*. New York: Chapman and Hall, pp. 97-107.

- Tilton, J. C., 2013. *RHSeg User's Manual: Including HSWO, HSeg, HSegExtract, HSegReader, HSegViewer and HSegLearn*, version 1.59: available via email request to James.C.Tilton@nasa.gov.
- Tilton, J. C., Y. Tarabalka, P. Montesano, and E. Gofman. 2012. Best merge region-growing segmentation with integrated nonadjacent region-object aggregation. *IEEE Trans. Geosci. Remote Sens.*, 50(11):4454-4467.
- Unsalan, C., and K. L. Boyer. 2005. A system to detect houses and residential street networks in multispectral satellite images. *Computer Vision and Image Understanding*, 98(3):423-461.
- Vincent, L., and P. Soille. 1991. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(6):583-598.
- Wang, C., N. Komodakis, and N. Paragios. 2013. Markov Random Field modeling, inference and learning in computer vision and image understanding: A survey. *Computer Vision and Image Understanding*, 117:1610-1627.
- Wang, Z. and R. Boesch. 2007. Color- and texture-based image segmentation for improved forest delineation. *IEEE Trans. Geosci. Remote Sens.*, 45(10):3055-3062.
- Williams, J. W. 1964. Heapsort. *Commun. ACM*, 7(12):347-348.
- Yang, L., Albregsten, F., Lonnestad, T. and Grottum, P., 1995. A Supervised Approach to the Evaluation of Image Segmentation Methods. In: *Proc. CAIP, Lect. Notes Comput. Sc. 970*, Prague, Czech Republic, pp. 759-765.
- Zahn, C. T., 1971. Graph-theoretic methods detecting and describing gestalt clusters. *IEEE Trans. on Computers*, C-20(1): 68-86.

- Zamalieva, D., S. Aksoy, J. C. Tilton. 2009. Finding compound structures in images using image segmentation and graph-based knowledge discovery. *Proc. of IGARSS*, Cape Town, South Africa.
- Zhang, Y. J., and J. J. Gerbrands., 1994. Objective and quantitative segmentation evaluation and comparison. *Signal Processing*, 39:3-54.
- Zucker, S. W., 1976. Region growing: childhood and adolescence. *Computer Graphics and Image Processing*, 5:382-399.