AN AUTOMATED DELINEATION TOOL FOR ASSISTED INTERPRETATION OF DIGITAL IMAGERY

Guillermo Castilla, Postdoctoral Fellow
Geoffrey G. Hay, Assistant Professor
Department of Geography, University of Calgary, 2500 University Dr. NW
Calgary, AB T2N 1N4, Canada
gcastill@ucalgary.ca

ABSTRACT

The manual delineation of vegetation patches or forest stands is a costly and critical stage in any landcover mapping project or forest inventory based upon photointerpretation. Recent computer techniques have eased the task of the interpreter; however a good deal of craftsmanship is still required in the delineation. In an effort to contribute to the automation of this process, we introduce *Size-Constrained Region Merging* (SCRM), a recently implemented software tool that provides the interpreter with an initial template of the to-be-mapped area that may reduce the manual digitization portion of the process. In essence, SCRM transforms an ortho-rectified aerial or satellite image (single or multi-channel), into a polygon vector layer that resembles the work of a human interpreter, whom without *a priori* knowledge of the scene, was given the task of partitioning the image into a number of homogeneous polygons all exceeding a minimum size (i.e., minimum mapping unit). We provide background information on SCRM foundations, workflow and usage, and illustrate its application on a Quickbird image acquired over a rural area.

INTRODUCTION

For decades, photointerpretation has been, and to a good extent is still, the method of choice for producing fine-scale forest and landcover maps. Recent computer techniques have eased the task of the interpreter, whom is now able to directly digitize on-screen without having to previously draw polygon outlines on hardcopy photos. More recently (circa 2000), innovation in the RS/GIS market with the *eCognition*® software (Definiens, 2007) provided users with the first commercial tools for *GEo-Object-Based Image Analysis*. GEOBIA* is aimed at replicating (and or exceeding experienced) human interpretation of RS images in automated/semi-automated ways (wiki.ucalgary.ca/page/GEOBIA). GEOBIA, which is inspired by the successful Object Oriented Design, uses geographic objects in addition to classes in order to model the landscape. The class relations among objects are represented in 'kind of' hierarchies (taxonomies) that provide inheritance; and structural relations among objects are represented in a 'part of' hierarchies (partonomies) that provide encapsulation (information hiding) (Booch, 1991). The first step in GEOBIA is image segmentation, the partitioning of the image into a set of jointly exhaustive, mutually disjoint regions that are more uniform within themselves than when compared to adjacent regions. These regions are subsequently used as basic units to form classified geo-objects. Unlike typical pixel-based classifications, an object-based classification takes into account not only conventional features such as radiometric signatures, but additional features that cannot be derived for individual pixels (e.g., size and shape features), and most importantly, relational features between regions. With such capabilities, GEOBIA has the potential to supersede not only conventional pixel-based methods, but also photointerpretation. These prospects are beginning to be recognised by leading companies in the sector of RS image analysis, who progressively are including segmentation modules in their recent product releases.

However, GEOBIA is still in its infancy. Image understanding is a complex cognitive process for which we may still lack key concepts. Since significant research remains until fully automated RS image interpretation is achieved, the general approach should be a pragmatic one. As such, the short term goal, rather than trying to replace human interpreters, would be to support them in generating more timely, consistent and accurate products (Leckie et al., 1998). Therefore, new and or better tools are required that produce incremental improvements in these areas. They need not provide final solutions or 100% correct results, they simply need to be tools that are useful and that can be

* GEOBIA is a recent sub-discipline of GIScience devoted to partitioning remote sensing (RS) imagery into meaningful image-objects, and assessing their characteristics through spatial, spectral and temporal scale. GEOBIA 2nd International Conference will be held in Calgary in August 2008 (www.ucalgary.ca/GEOBIA).

easily corrected when they go awry. Specifically, they must be simple to apply, not require expensive equipment, not substantially alter the mapping workflow, nor involve inordinate fine-tuning by the interpreter (Leckie et al., 1998). In order to facilitate these requirements, we introduce our *Size-Constrained Region Merging* (SCRM) tool.

Essentially, SCRM transforms an ortho-rectified aerial or satellite image (single or multi-channel), into a polygon vector layer. This layer resembles the work of a human interpreter, whom without *a priori* knowledge of the scene, was given the task of partitioning the image into a specific number of relatively homogeneous polygons all exceeding a minimum size. This layer may then be used as an initial template in the task of the interpreter, who simply needs to aggregate (and sometimes correct) pre-delineated regions by easy drag-and-click operations. We note that SCRM results are only meant to be an intermediate aid for the work of the interpreter, since i) SCRM only considers radiometric features for the segmentation, and ii) the correspondence between radiometric similarity and semantic similarity is not straightforward. We also note that although the SCRM sequence includes procedures other than region merging, we named the algorithm after this condition because it is the most influential step.

The objective of this paper is to provide an overview of SCRM as a tool for assisted photointerpretation. We i) provide conceptual foundations that underlie the SCRM approach to image segmentation; ii) briefly describe the SCRM workflow; iii) illustrate with an example how it may be used as an automated delineation aid for computer-assisted photointerpretation; and iv) provide some final remarks.

SCRM CONCEPTUAL FOUNDATIONS

The Nested Hierarchical Patch Paradigm

SCRM is a segmentation tool for GEOBIA, which is especially suited for implementing a new paradigm (Wu an Loucks, 1995; Woodcock and Harward, 1992) where the landscape is modelled as a hierarchical system having both vertical structure –composed of levels; and horizontal structure –consisting of patches or *holons* (wholes made of other wholes). The integration of each patch is such that the subunits compounding it (i.e. patches from the next lower level) interact more strongly or more frequently among them than with subunits of neighboring patches. Hence, higher levels are characterized by slower and larger entities loosely integrated (e.g. a forest), whereas lower levels are characterized by faster and smaller entities (e.g. a tree) more tightly integrated than the former. Therefore, at different levels, a patch may be an entity ranging from the area covered by an isolated tree to an island continent. Patches from different levels may overlap in size, but the mean size of patches within each level must increase steadily with the level. Since there are no crisp distinctions between adjacent levels in the hierarchy, some arbitrary size constraints must be established in order to formally define each level. SCRM takes those constraints as input parameters and constructs a partition of the imaged landscape that conforms to them.

Blobs: The Basic Perceptual Units

SCRM begins the region merging process with a fine partition made of primal regions or *blobs* derived from image morphology. A blob is a perceptual concept that refers to a tiny homogeneous region in the image, darker, brighter or of different hue than its surroundings. Being internally homogeneous and different from its exterior, a blob may be also viewed as the *basin of attraction* of a perceptual attractor, where all pixels within that basin may be perceived as forming a distinct single entity. Under this view, blobs are the basic perceptual units, just as pixels are the physical basic units. Perceptual attractors may be associated to local minima in a transformed image where the *digital number* (DN) of pixels is proportional to the difference with adjacent pixels in the original image. Thus, blobs can be formally defined as the basins of attraction of local minima in a gradient magnitude image derived from the original one through some dissimilarity metric. This definition enables us to establish a powerful and innovative link between catastrophe theory, graph theory and morphological image analysis (for additional details, see Castilla 2003).

Catastrophe Theory

According to René Thom's (1975) topological theory of attractors, also known as Catastrophe Theory, every object in an image can be represented as an attractor of a dynamical system on a space of internal variables. Thus, an object may be recognized only when the corresponding attractor is stable. Stability is attained by a process, called *morphogenesis*. This consists in the disappearance of the attractors representing the initial unstable forms, and their replacement (i.e., capture) by the attractors representing the final forms, which is the observable state of the object(s). An attractor can be thought of as the centroid of an object, so that pixels belonging to the object are more attracted to it than to the centroid of neighboring objects. However, in the boundaries between objects, this attractive force becomes unstable and an infinitesimal move in one or another direction may produce a change of attractor: these are the singular pixels, (i.e., edges or boundaries) that define the spatial structure of the image.

Graph Theory

We have translated this account into graph theory by considering the image as the initial state of a planar dynamic network consisting of triangular meshes made up of nodes (pixels) connected through links via which the nodes interact. This interaction consists of quantitative luminance exchanges between the nodes. The intensity of this interaction is regulated by proximity in feature space, decreasing rapidly with dissimilarity distance; and it is formalized through a weight allocated to each link. A link may be active, if there is some noticeable interaction through it, or inactive if its weight is nearly zero. To simulate morphogenesis we allow the network (i.e., image) to evolve through several cycles in which the state of a node is dependent on the state of adjacent nodes in the previous cycle, according to the above interaction mechanism. During this process, some nodes interact more strongly between themselves, while some others stop interaction. This induces a coherent behavior of nodes within local groups. Eventually some inactive links may become active, opening paths between nearby groups so that small groups are captured by larger ones. After a few cycles, the network reaches a steady state far from equilibrium (where equilibrium would represent a uniform distribution of luminance across the network, i.e. a flat image). Thus the network evolves towards a piecewise constant image in which the pixels within each local group have roughly the same value. Each local group has a node that acts as an attractor in feature space, i.e., a node whose basin of attraction is the local group itself. Then the remaining attractors within the network are the stable attractors, and their respective basins of attraction coincide with blobs. As a result, this evolution can be viewed as a self-organization of the image into perceptual units (Castilla, 2003). The practical implementation of this process is a non-linear diffusion filtering (see *Image smoothing* subsection).

Image Morphology

The set of singular pixels demarcating the stable basins of attraction is obtained via the watershed algorithm, a morphological segmentation method commonly used in Computer Vision and Biomedical Imaging (Meyer, 2001). This algorithm extracts the network of ridges (watersheds, or drainage divides) that exist in the input image (usually a gradient magnitude image) when it is considered as a Digital Elevation Model (DEM). In our method, since the filtered image (i.e., the final state of the above network) is a quasi-piecewise constant, its corresponding gradient magnitude image resembles a DEM from a lunar plain full of craters of different shape and size. Here, each crater is the basin of attraction of a local gradient minimum, which is located at the bottom of the crater. Now the concept of attractor becomes more evident by introducing an analogy with gravity. For example, if a basketball is dropped at any location within a crater, it will be 'attracted' towards the bottom and eventually come to rest there. The event of the ball getting trapped in some pit (i.e., an unstable attractor) near the bottom is precluded by filtering, which smoothes out the crater surface. However, at the very top of the crater, the attractive force exerted by the base of the crater becomes unstable, as any small shift may result in the ball falling in or out. Therefore ridges may be defined as the set of singular pixels bounding the basins of attraction, which conveniently also happen to be the pixels that the watershed algorithm foregrounds as boundaries of catchment basins. When blobs are viewed in this way, it becomes apparent that the watershed algorithm is more than just another segmentation method: it is the tool for applying Thom's (1988) *semiophysics* to image analysis (Castilla, 2003). The merit of Thom's work is that it provides a link between the quantitative physical theory of the world and our qualitative daily experience (Smith 1995). Our claim, which we believe is an important contribution to this field, is that this link can be realized via gradient watersheds when we observe the world through images.

From Blobs to Patches

The partition rendered by the watershed algorithm is usually too profuse to form a meaningful representation of the imaged landscape. Most defined regions (blobs) will be surrounded by neighbors of similar appearance, hence they can be aggregated into a single unit that is more likely to constitute a unitary patch. In addition, most blobs will have a size that is below the one required for a region to be considered a patch of the targeted hierarchical level of landscape analysis, i.e., the focal level. Therefore, the watershed partition has to be simplified according to a similarity criterion until the size constraints defining the focal level are met. After this, the remaining regions can be made semantically coherent (so that each region or segment can be assumed to correspond to a geo-object or patch) either via a semiautomated GEOBIA procedure, or through photointerpretation. In the latter case, the interpreter simply needs to aggregate adjacent regions that in her opinion are part of the same unit, or more rarely, split regions consisting of semantically different parts that were merged because of their similar appearance.

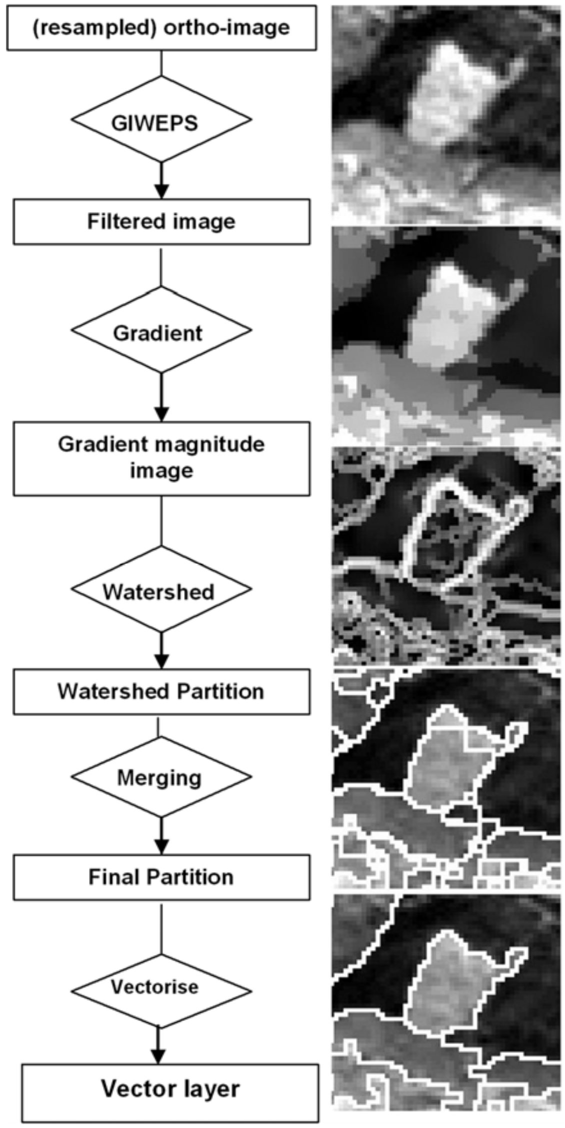

Figure 1. Overview of the SCRM workflow.

SCRM WORKFLOW

SCRM source code was written by Dr Castilla in IDL® (ITTVIS, 2007a), and can be run either within the commercial remote sensing software ENVI® (ITTVIS, 2007b) as a user extension, or stand-alone in conjunction with the IDL Virtual Machine. In order to use SCRM, four parameters must be specified: i) the *desired mean size* of output polygons (DMS -in hectares); ii) the minimum size required for polygons, or *minimum mapping unit* (MMU - in hectares); iii) the *maximum allowed size* (MAS -in hectares); and iv) the minimum distance between vertices in the vector layer, or *minimum vertex interval* (MVI -in metres). MVI is an indication of the positional accuracy of boundaries and is internally used to define the working pixel size (i.e., spatial resolution). The default values of these parameters are shown in Table 1. We note that there are also some optional parameters, like the possibility to take into account both the internal texture of regions and the saliency of edges separating regions. With the latter option, merging of similar regions separated by a strong edge (like a narrow road separating two fields with the same crop) is precluded.

Table 1. Default values of SCRM input parameters for some common multispectral imagery and suggested scale for the final map product.

Sensor	Minimum Mapping Unit (ha)	Desired Mean Size (ha)	Maximum Allowed Size (ha)	Minimum Vertex Interval (m)	Potential map scale
Quickbird MSS	0.01	0.06	0.30	5.0	1:10,000
SPOT MSS	2.50	10.00	50.00	20.0	1:50,000
ETM+ MSS	22.50	90.00	450.00	60.0	1:100,000

SCRM workflow is as follows (Fig. 1). The input image (previously ortho-rectified to some cartographic projection) is (if necessary) resampled to a suitable pixel size, and then filtered with *Gradient Inverse Weighed Edge Preserving Smoothing* (GIWEPS, Castilla, 2003). The output of this step is an almost piecewise constant image, from which the gradient magnitude is computed. This gradient magnitude image is then searched for local minima, and the area of influence of each minimum is contoured and labeled with the watershed algorithm. The resulting regions are aggregated iteratively by increasing dissimilarity until they all exceed the size of the minimum mapping unit (MMU). Then, the labeled image containing the final partition is converted into a vector layer.

Image Resampling

The default minimum vertex interval (MVI) is double the pixel size of the input image, since this is the spline interval that is applied to interpolate the centers of watershed pixels (see *Vectorization* subsection). If the final map product is intended to depict the theme of interest at a coarser scale, then segmentation should be performed at a resolution that matches this cartographic requirement, in the same way interpreters should not zoom beyond a certain visualization scale while digitizing arcs. On the one hand, too fine a scale often overwhelms the operator (be it man or machine) with excess detail and limited context. Conversely, as a consequence of the fractal nature of landscapes, the length of arcs increases indefinitely as the resolution increases. Thus, the finer the pixel size, the more intricate the arcs. Therefore, in this case the input image should be upscaled to a resolution balancing edge simplicity and accuracy. This is done by resampling the image to half MVI by simple pixel averaging. When the user is unsure about what minimum vertex interval fits their application, they may set MVI equal to the boundary positional accuracy required for the final map, as it will on average be better than this value.

Image Smoothing

A consequence of hierarchical patchiness in landscapes is the presence of areas of coarse texture in all types of remote sensing imagery, irrespective of spatial resolution. Coarse texture, or high local variance, is mainly due to the existence of recurrent elements that are large enough to produce some variation in the image, but too small to be resolved by the sensor, such as individual trees in a *Thematic Mapper* image. Since no information can be retrieved about their individual shape, they cannot be included individually in a shape-oriented representation of the scene derived from that image. Therefore they have to be captured within the larger element they are a part of (e.g., a forest stand). On the other hand, coarse texture areas are characterized by a high density of local luminance extrema that produce a considerable amount of spurious minima in the corresponding gradient magnitude image. Following the discussion in the background section, these minima may be regarded as perceptually unstable attractors, which we remove by processing the image with a non-linear diffusion filter that simulates the aforementioned morphogenesis. Non-linear diffusion is an iterative process where *diffusivity* (the rate of luminance exchange between adjacent pixels) changes according to the evolution of the *local gradient* (difference in brightness for single channel images, or dissimilarity for multichannel images between each pixel and their immediate neighbors). In our implementation, called Gradient Inverse Weighted Edge Preserving Smoothing (GIWEPS), the new DN of a given pixel is the weighted mean of the current DNs of its eight neighbors. The weight of each neighbor is proportional to its Euclidean distance (in the feature space) to the pixel under examination, and the proportion is governed by a heuristically defined diffusivity parameter. The filter is applied iteratively, until the difference between consecutive output images is negligible, which typically occurs in a few iterations. Further details can be found in Castilla (2003).

Gradient Magnitude Image

If one considers a given grey-level image as a Digital Elevation Model (DEM), then the gradient magnitude image is the slope map that corresponds to that DEM. Thus, at each pixel of a grey-level image, the gradient magnitude is the slope of the steepest descent line crossing that pixel. In the case of multiband images, the slope describes the variations in similarity of adjacent pixels across the image. The dissimilarity measure used here is the Euclidean distance between points (pixel signatures) in the feature space. At each pixel of the image, the gradient magnitude is computed as the square root of the sum of squared dissimilarity distances between the East and West neighbors, and the North and South neighbors respectively. Note that the gradient minima are those pixels whose value is lower than that of their eight neighbors in the gradient magnitude image.

Watershed Partition

The application of the topographic concept of watershed to the field of image analysis was introduced by Beucher and Lantejoul (1979) and later implemented into an efficient algorithm by Vincent and Soille (1991). In the SCRM workflow, the idea is to consider the gradient magnitude image as a DEM. The goal is to find the drainage divides, or watersheds, of that virtual territory, which is achieved by simulating a gradual immersion of the DEM. In the output partition, watershed (boundary) pixels are set to zero, whereas non-zero pixels have as DN the numeric label of the region (blob) to which they belong.

Region Merging

In this step the regions of the watershed partition (i.e., blobs) are aggregated until all regions in the partition are larger than the specified minimum size (MMU). The merging sequence is such that the homogeneity of the resulting regions is maximal given the size constraints. The dissimilarity measure used as merging criterion is the Euclidean distance in a feature space where each dimension corresponds to a channel of the image. Thus, the signature of a region, from which dissimilarity to its adjoining neighbors is computed, corresponds to the coordinates of the region centroid in the feature space. That is, a signature is an n-component vector where each component is the mean value in each of n channels of pixels belonging to the region. After a merge, the signature of a new region is the weighted (by size) mean of the signatures of the two merged regions. In this way, region signatures are computed from the original image only once, at the beginning of the procedure. The same can be said about the adjacency table (an array returning the list of neighbors of any given region), which is first computed from the watershed partition and then updated using Boolean algebra. From this adjacency table (AT) and the signature list (SL), the identification of the most similar neighbor (MSN) to each region is trivial. In each iteration, the two adjacent regions that merge are those best fitting, (i.e., those having the least dissimilarity distance from the set of candidate pairs). Next, the AT, SL and MSN arrays are updated, the new best fitting candidate pair is identified, and a new iteration proceeds. During the first several hundred iterations, the list of candidate pairs consists of every pair of adjacent regions in the image. As the merging proceeds, smooth low contrast areas become occupied by increasingly larger regions, and the maximum size constraint (MAS) comes into play. At a given iteration, if the best fitting pair consists of two regions both exceeding MAS, then it is not allowed to merge. Furthermore, this pair is permanently withdrawn from the candidate list, and the next best fitting pair is selected for merging. The merging continues this way until the sum of i) the number of regions currently larger than the minimum allowed size (MMU), plus ii) the expected number of final regions that may result from the area currently occupied by regions smaller than MMU, is less than the expected number of final regions (i.e. the image area divided by DMS). This is a partial stop criterion that guarantees that the final mean size of regions will be close to DMS. Thereafter, the candidate list is restricted only to those pairs where at least one of both regions is smaller than MMU. Once the merging is completed, a new image (the SCRM partition) is created from the watershed partition by replacing the DN of pixels inside each blob with the new label registered in the corresponding position of FLL. Finally, watershed pixels lying in the interior of final regions are filled with the numeric label of the corresponding region.

Vectorization

The last step in the workflow is to convert the SCRM partition into a polygon vector layer and save it as an ESRI shape-file (format .shp). In order to proceed, the centres of boundary (zero-valued) pixels are considered the initial vertices forming the vector layer. This is analogous to considering boundary pixels as a transition zone between patches that can be represented by its medial axis. Then nodes (junctions connecting arcs) are the centres of those boundary pixels having more than two non-diagonal neighbors that are boundary pixels. Arcs on the other hand correspond to chains of boundary pixels that start and end by a node. Finally, vector units (polygons) are delimited by the set of arcs bounding the corresponding region. In order to give a 'natural' appearance to arcs, a spline interpolation is applied to the centroid of each three consecutive vertices within the arc. The smoothed arc is further simplified with a proprietary implementation of the Douglas-Peucker (1973) algorithm, which deletes redundant vertices using a user-defined tolerance whose default value is half the pixel size (Fig. 2). These results are then saved as a shape file (ESRI format .shp), and the associate database file (.dbf) is filled with radiometric statistics about each polygon (i.e., min, max, mean DN and standard deviation).

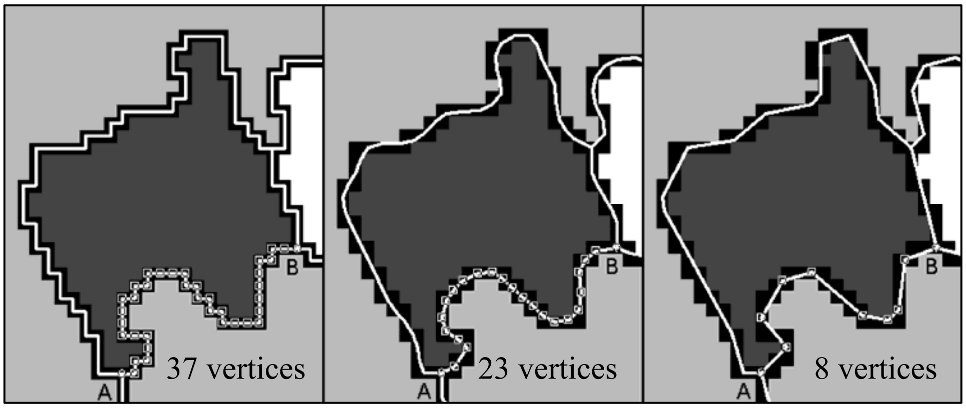

Figure 2. Vectorization sequence in a sample partition. Left: raw vector. Centre: vector after spline interpolation. Right: final vector after Douglas-Peucker simplification. The vertices of arc AB are highlighted.

EXAMPLE

In this section we illustrate how a SCRM output vector layer may be used as an initial template for computer assisted photointerpretation. Before proceeding, the user should have a 'feel' for how 'broken up' the scene needs to be. For example, if we are interested in delineating forest stands with a 10 ha average size, a suitable *Desired Mean Size* (DMS) would be some 2.5 ha. In this way we would generate sufficient units to avoid excessive manual digitization in a latter stage. The *Maximum Allowed Size* (MAS) could be set to 10 ha, so that units larger than 20 ha would be rare, or left blank, if this is not a concern. The *Minimum Mapping Unit* (MMU) of the final map must be known in advance, or the user must recognize that any region below the default size will not be retained in the output partition, no matter how distinct the region is. The last SCRM input parameter, the *Minimum Vertex Interval* (MVI), can also be set intuitively. If there is no formal requirement, a good rule-of-thumb is the recommended digitizing visualization scale. For example, for a visualization scale of 1:10,000, a reasonable MVI would be 5 m, and 50 m for 1:100,000.

Fig. 3 shows a 1.4x0.92 km^2 sub-scene from a semi-natural landscape centred on *Muskilda* hill, a 50 ha forest near Estella (Navarra, Spain), covered by *Quercus faginea* trees and shrubs, and surrounded by vineyards, pastures and ploughland. The color composite (RGB bands 4, 3, 2) is a 2.8 m pixel Quickbird-2 multispectral image acquired in September 2004. This is the base image we will use in figures 4-6 to illustrate how SCRM may be used for photointerpretation.

Fig. 4 shows SCRM results for DMS=1.5 ha, MMU=0.5 ha, and MVI=10 m (no MAS restriction). Imagine we want to compile a forest map with a minimum mapping unit of 2 ha. The partition produced from Fig 7 SCRM-inputs would be too profuse for this purpose. In addition, it has units, like the 1.4 ha rectangular patch of scrub to the left of the forest (marked with an 'A'), that even if they were densely populated by trees, would not qualify as forest in this map. Fig. 5 illustrates SCRM results for DMS=6 ha, MMU=2 ha, and MVI=10 m (no MAS restriction). Here, the aforementioned scrub no longer constitutes a separate unit, as it is smaller than the MMU. Also, some regions now form part of a different aggregate than in the previous partition, like the small area marked with an 'X' in the right side of the forest in Figs. 4 and 5. In Fig. 4, this region had as its most similar neighbor, the scrub located below it. However, in Fig. 5, the bright barren area surrounding it has previously merged with some slightly darker agricultural fields, lowering the average brightness of the aggregate, while at the same time the scrub has merged with the (also darker) forest, so that now the most similar neighbor is the barren area instead of the scrub.

Figure 3. A 1.4x0.92 km^2 Quickbird (02) RGB 432 sub-scene showing Muskilda hill, a 50 ha oak forest to be delineated.

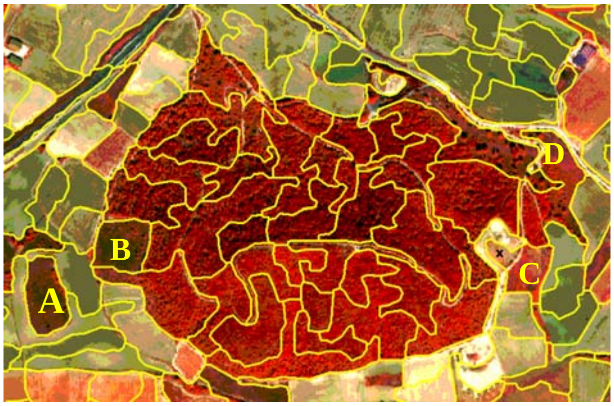

Figure 4. SCRM results applied to Muskilda hill (Fig. 6) using DMS=1.5 ha, MMU=0.5 ha, and MVI=10 m.

Figure 5. SCRM result used as template for the example in Fig.9 (DMS=6 ha, MMU=2 ha, and MVI=5 m).

The partition of Fig. 5 is better suited than that of Fig. 4 as an initial template to delineate our forest map, as it has no unit smaller than the specified 2ha MMU. Therefore, the basic operation is to manually merge connected regions that in our opinion are forests. In this example, this would be done in a few seconds (in any GIS with editing capabilities) by hold-clicking and dragging with the left mouse button over the polygons we want to merge as to select them, and then right clicking to confirm the merge. The rectangle shown in Fig. 5 represents an instance of such movement that would produce the merging of the eight polygons within Muskilda hill. Next, we would need to correct several small areas along the perimeter of the newly formed polygon that look similar to the forest but that are actually scrub with less trees than would qualify as forest (Fig. 6). This is the case of the two patches lying at both extremes of the lower half of Muskilda, and also of its NE corner (marked respectively 'B','C' and 'D'). Finally, there is a gravel pit surrounded by a thin corridor of trees in the upper right half of Muskilda that we have decided not to retain within the forest because of the narrowness of the corridor. All these operations can be done with a stream digitizing tool that manually splits the undesired parts into separate polygons, where the latter would be subsequently merged to the surroundings using the above standard procedure. Fig. 6 illustrates how Muskilda forest would look in the final map. In this example, less than ten percent of the polygon (those arc segments that appear labeled) has been delineated manually.

Figure 6. Final delineation of Muskilda forest.
Only arc segments AB, CD, EF and GH have been digitized manually.

FINAL REMARKS

In this paper we have introduced Size Constrained Region Merging (SCRM), a novel segmentation method that may be used for computer assisted photointerpretation. SCRM transforms a single or multichannel ortho-image into a polygon vector layer (.shp) that may be used by the interpreter as an initial template. SCRM may be applied to any kind of RS imagery (in geotiff, jpeg, or ENVI format), of any size (images larger than 2 Mpixel at working resolution are subject to tiled processing) and any number of channels. Typical processing time (using an Intel Pentium IV, 2.3 GHZ, with 1 GB of RAM) is less than two minutes per Mpixel. After having applied SCRM to many images, none of the output partitions produce a visual impression of a 'bad segmented' image. Furthermore, compared to other segmentation algorithms embedded in already available commercial software, SCRM is less demanding computationally; requires only intuitive input parameters that enable the user to explicitly control the level of cartographic generalization (both the size distribution of polygons and the edge complexity) applied to the image; and tackles the fractal nature of landscape and its hierarchical structure in a conceptually coherent manner. Moreover, SCRM is grounded on a solid conceptual basis (Castilla 2003), an asset that many segmentation algorithms lack. Notwithstanding, a thorough comparison with other segmentation methods is desirable in order to fully evaluate the pros and cons of our procedure. In particular, the usability of SCRM vectors – with regard to time savings during the interpretation process - needs to be confirmed by empirical studies. In this respect, a test carried out by Castilla (2006) to evaluate the possibility of updating the Forest Map of Spain with SCRM yielded a promising result, since more than 80% of the edges of the final map came from the automatic segmentation, which may represent a 40% time-saving in the photointerpretation.

ACKNOWLEDGMENTS

The authors wish to thank Geosys (www.geosys.es) for providing them with the satellite image used in the example. This research has been generously supported in part by a University of Calgary Start-up Grant, and an Alberta Ingenuity New Faculty Grant to Dr Hay. The opinions expressed here are those of the Authors, and do not necessarily reflect the views of their funding agencies. NB. This contribution is a reduced and updated version of a paper accepted in PE&RS (Castilla et al., 2007), to appear this summer.

REFERENCES

Beucher, S. and C. Lantuejoul, (1979). Use of watersheds in contour detection. *Proceedings of the International Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation*, September 1979, Rennes, France.

Booch, G., (1991). *Object-Oriented Design with Applications*. The Benjamin/Cummings Publishing Company Inc., Redwood City.

Castilla, G., G. J. Hay and J. R. Ruiz, (2007). Size-constrained region merging (SCRM): an automated delineation tool for assisted photointerpretation. *Photogrammetric Engineering & Remote Sensing* (in press).

Castilla, G., (2006). Actualización semiautomática de la malla de recintos del Mapa Forestal de España. *Investigación Agraria. Sistemas y Recursos Forestales*, December Special Issue: 14-23.

Castilla, G., (2003). *Object-oriented Analysis of Remote Sensing Images for Land Cover Mapping: Conceptual Foundations and a Segmentation Method to Derive a Baseline Partition for Classification*. Ph.D. Thesis, Polytechnic University of Madrid, URL: http://www.montes.upm.es/servicios/biblioteca/tesis/GCastillaTD_Montes.pdf, Madrid, Spain.

Definiens, (2007). Definiens Professional 5.0. Definiens AG | Geospatial Solutions München, Germany. URL: http://www.definiens.com/products/definiens_professional.php, (last date accessed: 1 February 2007).

Douglas, D. and T. Peucker, 1973. Algorithms for the reduction of the number of points required to represent a line or its caricature. *The Canadian Cartographer*, 10 (2): 112-122.

ITTVIS, (2007a). IDL The Data Visualization & Analysis Platform. ITTVIS, Boulder, Colorado, USA. URL: http://www.ittvis.com/idl/, (last date accessed: 1 February 2007).

ITTVIS, (2007b). ENVI - The Remote Sensing Exploitation Platform. ITTVIS, Boulder, Colorado, USA. URL: http://www.ittvis.com/envi/, (last date accessed: 1 February 2007).

Leckie D. G., M.D. Gillis, F. Gougeon, M. Lodin, J. Wakelin and X. Yuan, (1998). Computer-assisted photointerpretation aids to forest inventory mapping: Some possible approaches. *Proceedings of the International Forum on Automated interpretation of High spatial resolution digital imagery for forestry*, 10-12 February 1998, Victoria, British Columbia (Canadian Forest Service, Pacific Forestry Center, Victoria, B.C.), pp. 335-343.

Meyer, F., (2001). An overview of morphological segmentation. *International Journal of Pattern Recognition and Artificial Intelligence*, 15 (7): 1089-1118.

Smith, B., (1995). The structures of common-sense world. *Acta Philosophica Fennica*, 58: 290-317.

Thom, R., (1975). *Structural Stability and Morphogenesis: An Outline of a General Theory of Models*. Benjamin-Cummings Publishing.

Thom, R., (1988). *Esquisse d'une sémiophysique*. Paris: InterEditions.

Vincent, L. and P. Soille, (1991). Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13 (6): 583-598.

Woodcock, C.E. and V.J. Harward, (1992). Nested-hierarchical scene models and image segmentation. *International Journal of Remote Sensing*, 13 (16): 3167-3187.

Wu, J. and O.L. Loucks, (1995). From balance-of-nature to hierarchical patch dynamics: a paradigm shift in ecology. *Quarterly Review of Biology*, 70: 439-466.