LA-UR- 08-7898

Approved for public release; distribution is unlimited.

Title: Hierarchical image feature extraction by an irregular pyramid of polygonal partitions

Author(s): Skurikhin, A.N.

Intended for: ASPRS 2009 Annual Conference, Baltimore, MD, March 2009

HIERARCHICAL IMAGE FEATURE EXTRACTION BY AN IRREGULAR PYRAMID OF POLYGONAL PARTITIONS

Alexei N. Skurikhin, Technical Staff Member
MS D436, Space and Remote Sensing Group
Los Alamos National Laboratory
Los Alamos, NM 87545
alexei@lanl.gov

ABSTRACT

We present an algorithmic framework for hierarchical image segmentation and feature extraction. We build a successive fine-to-coarse hierarchy of irregular polygonal partitions of the original image. This multiscale hierarchy forms the basis for object-based image analysis. The framework incorporates the Gestalt principles of visual perception, such as proximity and closure, and exploits spectral and textural similarities of polygonal partitions, while iteratively grouping them until dissimilarity criteria are exceeded. Seed polygons are built upon a triangular mesh composed of irregular sized triangles, whose spatial arrangement is adapted to the image content. This is achieved by building the triangular mesh on the top of detected spectral discontinuities (such as edges), which form a network of constraints for the Delaunay triangulation. The image is then represented as a spatial network in the form of a graph with vertices corresponding to the polygonal partitions and edges reflecting their relations. The iterative agglomeration of partitions into object-oriented segments is formulated as Minimum Spanning Tree (MST) construction. An important characteristic of the approach is that the agglomeration of polygonal partitions is constrained by the detected edges; thus the shapes of agglomerated segments are more likely to correspond to the outlines of real-world objects. The constructed partitions and their spatial relations are characterized using spectral, textural and structural features based on proximity graphs. The framework allows searching for object-oriented features of interest across multiple levels of details of the built hierarchy and can be generalized to the multi-criteria MST to account for multiple criteria important for an application.

INTRODUCTION

The need for prior object-oriented segmentation is a central problem for pattern recognition. Methods which are able to parse images into meaningful object-oriented image chunks are therefore required for the successful automated image interpretation, content-based image retrieval, and object recognition. While low-level features can be easily extracted from images, they are not completely descriptive for image interpretation, as object is often not represented in single group of pixels. The disconnect between the low-level features and the high-level semantics required for image interpretation is known as "semantic gap" (Smeulders, et al., 2000). Bridging the semantic gap is difficult due to the high variability of objects' appearances, occlusions, complex spatial relationships between the objects, and possible concealment efforts to suppress objects' signatures. Higher spatial resolution of satellite images adds a complexity for automated extraction of image semantics. Increases in spatial resolution extend variation of objects' spectral, structural and textural characteristics. Existing techniques used to process lower resolution satellite images do not generalize well to higher resolutions, due to the changed appearance of features of interest (for a review, see Baltsavias, 2004; Schiewe, et al., 2001). For instance, roads become less line-like and more polygon-like, while differentiation of roads and rooftops based on spectral features is problematic, due to the fact that they can be made of similar materials (asphalt) and thus have similar spectral and textural properties (Herold et al., 2003a, 2004). Additionally, an object can always be represented in more than one way, and using more than one level of details (LODs). Therefore agglomeration of pixels into groups and then into larger object-oriented chunks takes into account contextual information in the form of spatial relations between the object-oriented chunks both on single LOD and across multiple LODs. Exploiting contextual information improves object classification by resolving ambiguities, when objects are similar based on features of individual objects, such as spectral characteristics, area, shape (Barnsley, et al., 2000; Blaschke, et al., 2000; Bock, et al., 2000; Fuller, et al., 2004; Strat, et al., 1991). Complementary, it was shown that the localization and recognition of an object of interest can be improved through the detection of non-object low-level features whose presence in the image background is correlated with the presence of the object of interest. For instance, (Torralba, 2003) exploits the spatial frequency characteristics of the image to detect most likely object locations in general outdoor images. To what extent the last approach can be efficiently generalized to satellite imagery remains an open research question, as the background in satellite images can vary a lot thus reducing correlation between the presence of specific low-level features and the presence of objects of interest. Additionally, the objects themselves are not necessary isolated images chunks; they are networks of spatially distributed structures of varying sizes (e.g. an industrial site). Therefore robust hierarchic object-oriented segmentation will be essential for a robust satellite image interpretation, map updating, and change analysis (Blaschke, et al., 2006; Hoffmann, et al., 2001; Liu et al., 2008; Walter, 2004; Zhan, et al., 2002).

This paper presents a framework for hierarchical image segmentation that builds upon the detection of spectral discontinuities, edges. The framework constructs an irregular pyramid that contains a stack of vectorized images of successively reduced LODs. The vectorized images are represented as sets of polygonized pixel patches (polygons) attributed with spectral, textural, and structural characteristics. We use a combination of constrained Delaunay triangulation and the Gestalt principles of visual perception, such as proximity and closure, to exploit structural relations between the detected edges, and build an initial set of polygons (seeds). A polygonized image is then represented as a graph with vertices corresponding to the polygons and edges reflecting polygon relations. This is followed by the iterative graph contraction using Boruvka's Minimum Spanning Tree construction algorithm. The graph contractions merge the polygonal patches until dissimilarity criteria are exceeded. The contribution of our work is the following. We introduce an approach to construct an irregular image pyramid based on triangular and polygonal tessellations adapted to the image content. Our adaptive polygon-based hierarchic image segmentation iteratively constructs image chunks on coarser LODs from the chunks built at finer LODs. This results in the image multi-scale spatial network model, that is used for object-oriented image analysis, including analysis of contextual relations. In addition, because number of the polygonized image partitions is much smaller than that of the image pixels, our approach significantly reduces computational complexity compared to conventional graph partitioning methods that are directly applied to the image pixels.

In the next section we present related research in the areas of object-oriented image analysis, and hierarchical image segmentation using graphs. Then, we present our framework for feature extraction and image segmentation. We describe methods to extract spatial neighbor relations, and MST-based hierarchical image segmentation. This will be followed by outlining the results obtained, and finally, last section presents the conclusions.

RELATED WORK

Object-Oriented Image Analysis

The object-oriented image processing relies heavily on successful image segmentation. However, while image segmentation based on color, texture, structural features, and multiresolution hierarchy is a very important step in the overall process of image analysis, it alone can not address all the problems of the automated image interpretation. The high-level semantics required for image interpretation is mostly contained in the objects' spatial relations. Thus, the segmented objects are input for higher-level object-oriented image analysis with the goal of extracting image semantics. The extracted information about the objects' relations can then be exploited with semantic networks that formalize knowledge about the scene. For instance, semantic networks have been used in aerial image analysis (Quint, et al., 1995), and for updating the maps in GIS database (Kunz, et al. 1997).

The exploitation of spatial relations between the detected objects leads to the detection of geometric arrangements of objects. Example sinclude building alignment, settlement partitioning. Spatial analysis facilitates the discrimination of classes with similar spectral responses through the detection of the relevant contexts for the classes; it is also useful for multitemporal reasoning, analyzing both the former classification of the detected object-oriented image partitions and the plausible class transitions in that particular time interval and location. This analysis can also be used to improve segmentation of individual objects constituting the constructed cluster. The studies of (Barnsley, et.al, 2000, Herold, et.al, 2003b; Lowell, et al. 1992; Mustière, et al., 2002) show that the spatial analysis and spatial metrics are indispensable for satellite image interpretation, improving landscape classification, and cartographic generalization.

Delaunay triangulation and related proximity graphs such as MST, Relative Neighborhood Graph, Gabriel Graph (Toussaint, 1980; Jaromczyk, et al., 1992) have been widely used in spatial analysis and spatial modeling. Much of the work has been devoted to the discovery of building structures and road patterns in urban settlements. (Anders, 2003) uses graph-based approach to detect "natural" groups of buildings by analyzing and removing graph edges linking individual objects. (Regnauld, 2001) applies a proximity graph based approach to recognize building groups. The proximity graph is segmented using criteria inspired by the Gestalt psychology (Wertheimer, 1958). These criteria (e.g. proximity, similarity, good continuation, common regions) specify the formation of perceptually significant and visually attractive patterns (Boyer, et al., 2000). The challenge with using these criteria is that by themselves they provide no general-purpose scheme to resolve potentially conflicting outcomes of their application into an overall satisfactory result. Probably one of the first applications of some of these principles, specifically proximity, to the graph based analysis of points was done by (Zahn, 1971); in satellite imagery points would correspond either to detected buildings or other objects. Further example of spatial analysis include land use classification in (Zhang, et al., 2002) that is based on Delaunay triangulation that derives spatial relations between the detected image objects. (Taubenböck, et al., 2006) perform urban classification using a combination of color, shape and neighborhood related features. (Sharma, et al., 2008) uses the Delunay and the Voronoi graphs to extract features characterizing road and hydrographic networks. (Skourikhine, 2006) uses Delaunay triangulation and Euclidian MST to reconstruct road network using a set of pixel patches spectrally pre-classified as candidate road fragments.

Less work has addressed the problem of identifying salient ("interesting") structures across varying backgrounds. The saliency is given by semantics. e.g. industrial-like cluster of buildings in remote area. Salient structures can be either single objects or groups of objects (e.g. industrial facility). The outstanding problem is to detect such structures as salient clusters, while generalizing other objects into larger geospatial groups. It is necessary to efficiently map image partitions across multiple scales of the built image hierarchy, while extracting and exploiting their spatial relations and their spatial context. This requires the development of methods to construct a fine-to-coarse hierarchical relations between graph-based spatial networks of objects, and to detect salient object-based clusters (e.g. represented as regions of interest) that can be further utilized either by an analyst or some other kind of high-level analysis (e.g. semantic networks). More research is also called for the usage of proximity graphs for map generalization (Anders, 2003), as well as for initial object-oriented segmentation that provides seed partitions, which the image hierarchy is built upon.

Finally, quality of segmentation is of crucial importance. In spite of significant advances in image segmentation, evaluation of segmentation techniques is difficult. This is due to lack of unique ground-truth segmentation of an image against which the output of an algorithm is compared. There are some approaches of how to measure quality for segmentation in general such as (Unnikrishnan, et el., 2007; Zhang, 1996). However, these approaches have not been originally developed for the field of remote sensing and need to be adjusted. Recent studies of (Corcoran, et al., 2007; Neubert, et al., 2004, 2006, 2008a, 2008b) provide an overview of possible techniques to evaluate a quality of segmentation of satellite images, and show some of the evaluation approaches. (Corcoran, et al., 2007) distinguishes supervised (when user's outlines are used) and unsupervised evaluation and proposes to develop a cost function for the evaluation of quality of unsupervised segmentation. (Neubert, et al., 2004, 2006, 2008a, 2008b) has performed an extensive testing of different image processing software and algorithms, using a variety of measures e.g. overlapping areas and the geometrical correctness of the segmented outlines. If segmentation produces a hierarchy of image representations, then different objects are segmented out on different LODs of the image hierarchy. As the hierarchy is built some of the detected objects can be generalized into higher-level structures on coarser LODs of the hierarchy. This raises the question of how to take into account this multiplicity of possible image interpretations.

Graph Based Image Segmentation

According to the graph-theoretic approach to the image analysis, image elements such as pixels, regions, or edges are described using a weighted graph; and image segmentation is formulated as a graph-partitioning problem. Typically segmentation is achieved by optimizing a global criterion that takes into account the similarity between possible image partitions relative to the similarity within each partition. The similarity is usually computed based on color, texture, and spatial neighbor characteristics of the corresponding image elements. Many graph partitioning based image segmentation algorithms have been developed (e.g., Boykov et al., 2001; Fowlkes, et al., 2003; Shi, et al., 2000; Weiss, 1999; Wu, et al., 1993). They use partitioning criteria such as the minimum cut (Wu, et al., 1993), the normalized cut (Shi, et al., 2000). These algorithms are related to spectral methods which find good partitions via the eigenvectors of a matrix derived from the graph affinity matrix. The eigenvectors computations incur high complexity that is of $O(N^2)$, or $O(N^{3/2})$ when exploiting the sparseness of the graph, where N is the number of data elements, pixels. This complexity makes spectral methods of graph partitioning unsuitable for segmentation of very large satellite images. A possible solution to this problem is either to reduce number of pixels that has to be processed or to replace pixels of the original image with higher level structures, thus reducing the number of data elements to process. Another problem is related to the use of global optimization. Because there is no single optimal scale for segmenting all the objects from satellite image simultaneously, segmentation algorithm should produce a hierarchy of LODs (scales), from which objects can be detected. For example, (Sharon, et al., 2006) have proposed a hierarchical image segmentation approach using an algebraic multigrid, starting on a select subset of image pixels. Naturally, the overall segmentation performance of such approaches depends on the strategy of selection of the seed pixels or regions.

We emphasize another category of algorithms that seek optimal image hierarchic partitioning through a sequence of local computations based on proximity graphs, specifically MST. In contrast with global optimization approaches, the MST based image segmentation seeks image partitioning by iteratively linking image elements through the lowest cost tree edges, which represent similarity of neighboring elements. One of the earliest applications of tree-based data clustering to visual like point data sets analyzed histogram of MST edges and investigated tree characteristics such as MST "relative compactness", tree diameter, and point densities (Zahn, 1971). In (Horovitz, et al., 1976) the tree-based concept was applied to image segmentation. It was suggested to use global homogeneity criterion to control construction of an irregular pyramid starting from a regularly sampled pixel grid. (Montanvert, et al., 1991; Jolion, et al., 1992) use irregular tessellations to generate an adaptive multi-scale image representation. The approach employs an irregular sampling of the pixel grid to build the initial (lower scale) image representations. The irregular sampling hierarchy is then recursively built from the lower scales. The result depends on the stochastic nature of the sampling procedure. (Xu, et al., 1997) uses Kruskal's algorithm to construct MST (Kruskal, 1956) of the image from a regularly sampled pixel grid. The tree is then partitioned by an optimization algorithm into subtrees based on the subtrees' spectral similarities. A set of produced subtrees represents a sought image partition. Similar to (Xu, et al., 1997), (Felzenszwalb, et al., 1998) starts from a regular pixel grid and uses Kruskal's algorithm to construct MST of the image. However, MST construction is based on thresholding a ratio of the variation between neighboring pixel patches and the variation within the patches. To avoid over-fragmentation (generation of too many small regions), the approach adjusts the measure of variation using the sizes of patches. The extent of this adjustment controls how easily small patches are agglomerated in comparison with the larger ones. While it works in many situations, nontrivial optimization of this size based term is required for satellite image segmentation, where image elements of interest (e.g, fences, pipelines) can be of small size (width) and may not stand out strongly of the image background. The approach (Haxhimusa, et al., 2004, 2006; Kropatsch, et al., 2007) is similar to (Felzenszwalb, et al., 1998) in how it controls the grouping of pixels into patches based on image variation. The approach (Haxhimusa, et al., 2004, 2006; Kropatsch, et al., 2007) uses Boruvka's MST construction algorithm (Borůvka, 1926) instead of Kruskal's algorithm that is used in (Felzenszwalb, et al., 1998). Computational complexity of Kruskal's algorithm for computing the MST of a graph is $O(ElogE)$, and the complexity of Boruvka's algorithm is $O(ElogN)$, where E is the number of edges in the graph. In contrast with Kruskal's and Prim's (Prim, 1957) MST construction algorithms, that build the MST one edge at a time, Boruvka's algorithm adds several MST edges at each stage (Fig. 1). In the context of image segmentation this characteristic of Boruvka's algorithm provides an efficient approach for simultaneous agglomeration of image elements into higher level structures; it also makes possible to introduce user priors to control the MST construction.

Our framework belongs to the image segmentation approaches, producing an irregular image pyramids. However, in contrast with the stated approaches that start from regular or irregular pixel grids we build an irregular hierarchy of image partitions starting from triangular and polygonal tessellations of the image. The irregular polygon based image segmentation hierarchy is iteratively built bottom-up using Boruvka's MST algorithm. LODs of the constructed pyramid contain graphs of the agglomerated objects, and spatial analysis based on proximity graphs (such as Delaunay triangulation network and it's related neighborhood graphs) can now be applied to extract and exploit contextual relations between the objects across multiple LODs.

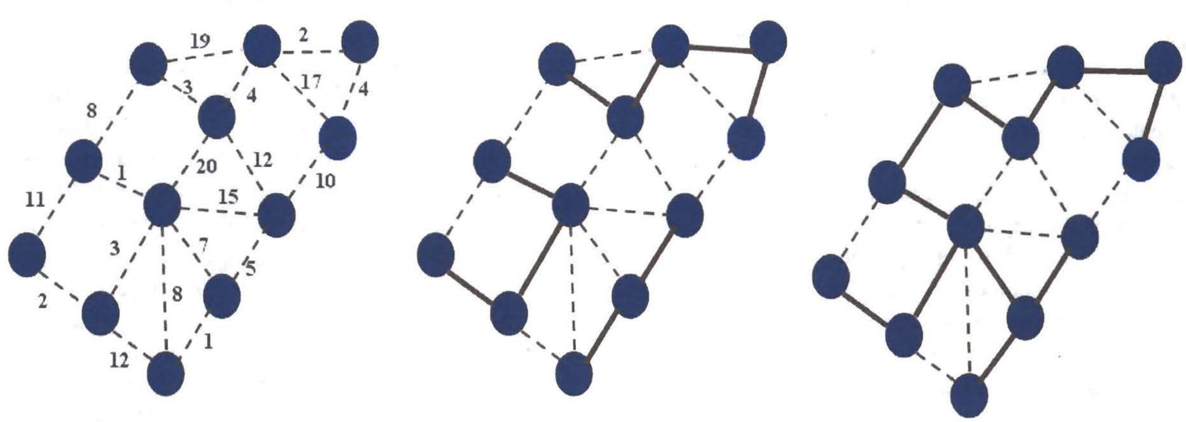

Figure 1. Iterations of Boruvka's MST reconstruction algorithm (from left to right). MST edges are shown in solid black lines; numbers next to the lines represent similarities between the nodes, the lower the number, the more similar the nodes are.

HIERARCHICAL IMAGE SEGMENTATION AND FEATURE EXTRACTION

In a polygon-based image pyramid, each level represents a polygonal tessellation of the image. The pyramid is built iteratively from bottom-up using only local interactions of the neighboring polygons. On the lowest level ($l=0$, fine level of detail) of the pyramid the polygons are constructed from an irregular triangular tessellation of the image; they are unions of triangles. On higher level ($l>0$, coarser level of detail) of the pyramid the polygons are unions of neighboring polygons on a lower finer level ($l-1$). The polygons on level l of the pyramid are considered as the vertices of an undirected graph G_l. The edges of the graph describe the adjacency relations between the polygons on level l. Thus $G_l =(V_l, E_l)$, where V_l is the set of vertices, and E_l is the set of edges. The derivation of G_{l+1} from G_l is formulated as construction of an MST of G_l. The built pyramid P is described as a set of graphs G_l representing the image in a fine-to-coarse hierarchy.

Extraction of Spatial Neighborhood

Spatial neighbors are usually defined as objects with common boundaries. However, evaluation of neighborhood relations and detection of spatial patterns require to consider more than nearest neighbors sharing common boundaries. It is necessary to take into account relations with more distant neighbors, *k*-order neighbors. Two objects are said to be *k*-order neighbors if the minimum number of regions to go through in order to get from one to another is ($k-1$). For an overview, see (Chen, et al., 2004). (Chen, et al., 2004) also develops a method to extract *k*-order neighbor relations based on the Voronoi diagram.

There are two general approaches to define spatial neighborhood: geometry based evaluating distance between objects, and topology based extracting spatial adjacency pattern between objects (e.g. a path). For a reason of computational efficiency we use the concept of topological neighborhood. We first introduce the maximal order of the topological neighborhood, k_{max}, that will be extracted and used in further analysis. Topological neighborhoods of individual polygons are then extracted using graph search algorithms such as depth-first search (DFS) or breadth-first search (BFS), which use k_{max} as the threshold defining the search region. Computational complexity of these search algorithms is $O(V^2)$ with adjacency matrix or $O(V+E)$ with adjacency list data representation, where V is number of polygons and E is number of edges linking the polygons. The polygon's topological neighborhood containing *k*-order neighboring polygons may then be attributed with pairwise polygon distances or other information. An example of the extracted topological neighborhood is shown in Fig. 2.

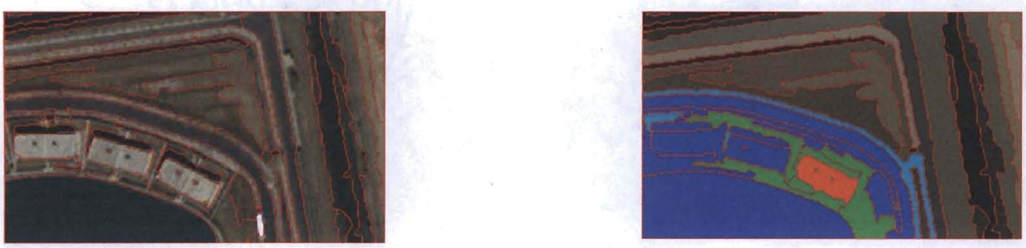

Figure 2. Extraction of multi-order topological neighborhood. Left: Contours (in red) of the generated polygons superimposed on top of the original image. Right: k-order neighbors of the red polygon: 1st order neighbors are in green, 2nd order neighbors are in blue, 3rd order neighbors are in cyan, others are shown with the computed color.

Construction of Seed Polygons

Polygons on the lowest (fine) level of a pyramid are built upon the triangular tessellation of the image. We employ the image vectorization approach (Prasad, et al., 2006a, 2006b) to process the generated triangle grid. First, we detect edges in the image, e.g using Canny edge detector (Canny, 1986) (Figs. 3a, 3b). This is followed by constrained Delaunay triangulation (CDT) (Schewchuk, 1996) where the detected edges are used as constraints for the triangulation (Fig. 3c). Thus, the CDT tessellation grid is adapted to the image content, since triangle vertices and edges reflect the structure and spatial adjacency of the detected edges, such as Canny edges. CDT generated triangular mesh is then processed by edge filtering. The filtering keeps constraints (such as detected Canny edges) and selectively deletes generated triangle edges. Triangle edge filtering is based on prespecified rules inspired by the principles of visual perception, such as proximity and closure. Proximity filters out triangle edges based on their length (Fig. 3e). As a result, the detected edges that are spatially close to each other are linked by the kept triangle edge. Otherwise, the detected edges remain separated. The closure rule is responsible for filtering out triangle edges which are bounded by the same detected edge (e.g., "U"-shape) or the same pair of detected edges (e.g., "| |" configuration) (Fig. 3f). This results in a set

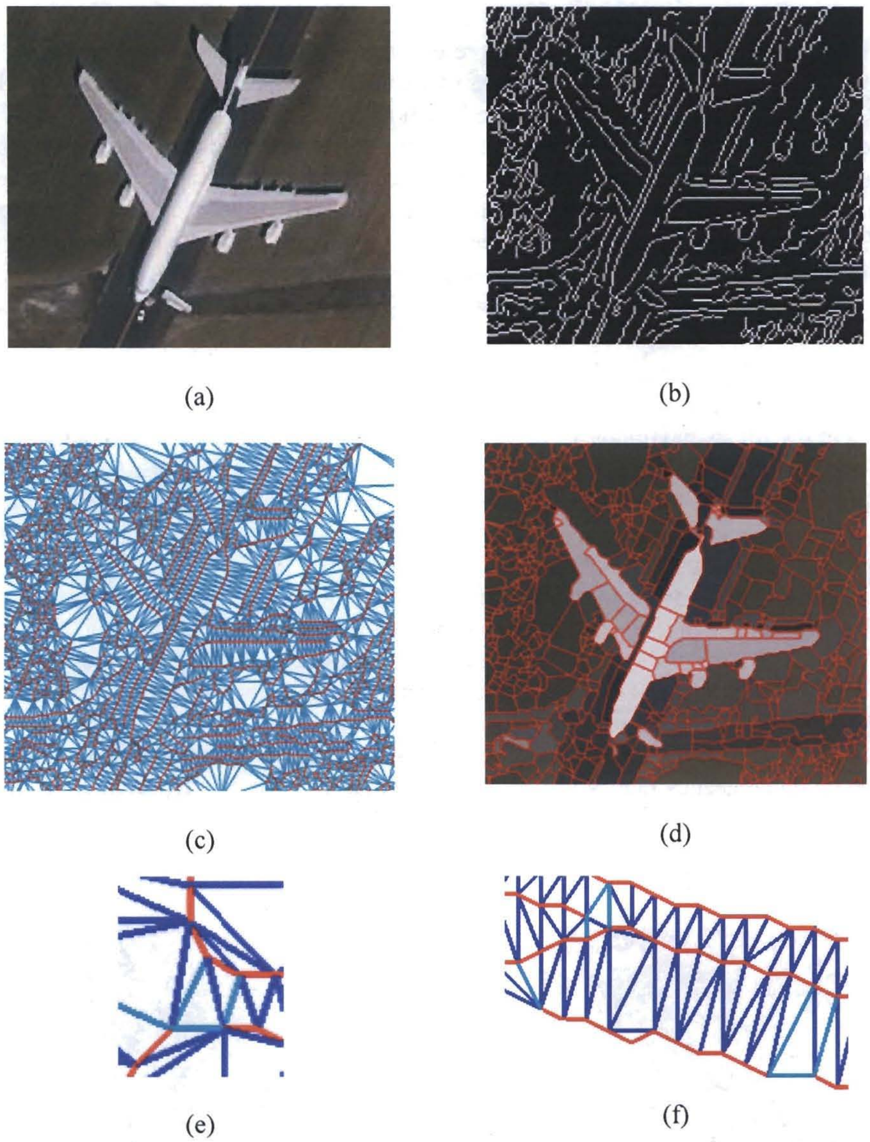

Figure 3. Construction of a fine LOD polygon-based image representation. We show a sequence of steps, which produce image segmentation on fine level of detail. (a) an original image. (b) edge detection using Canny edge detector. (c) generation of an irregular triangular mesh based on Constrained Delaunay triangulation (Canny edges are shown in red color, generated triangle edges are shown in cyan color). (d) contours (in red) of the created polygons superimposed on top of the original image. (e) an example of triangle edge filtering based on proximity. (f) an example of triangle edge filtering based on closure. Canny edges are shown in red color, deleted triangle edges are shown in blue, and the kept triangle edges (closing the gaps between Canny edges) are shown in cyan color.

of closed contours consisting of combination of the generated triangle edges and spectrally detected edges. Finally, a graph traversal algorithm (e.g., depth-first search or breadth-first search) groups triangles within the constructed closed contours into polygons (Fig. 1d). These polygons are assigned median color based on a sampling of pixels covered by the grouped triangles. Thus the image is segmented in a set of spectrally attributed polygonal patches (Fig. 2). The constructed polygon boundaries are built upon the detected edges, and thus reflect important spectral discontinuities in the image. This produces visually appealing results and reduces the amount of data, number of pixels to number of generated polygons, by 20-80 times depending on the image content. However, because the result is still over-fragmented, it is necessary to agglomerate the produced partitions into larger ones in order to perform object-oriented analysis.

Hierarchical Image Segmentation

While fine level-of-detail polygon-based image representation is over-fragmented, building larger polygons on top of it has the following distinct advantage: agglomeration of polygons will be implicitly directed in the sense that boundaries of agglomerated polygons will also be authentic to the image spectral discontinuities. The latter distinguishes our approach from other approaches, where selection of good seed pixels or pixel patches is challenging. It is due to the fact that pixel itself does not carry any object-oriented information. We extend our previous work (Skurikhin, 2008; Skurikhin, et al., 2008) by taking into account variogram-based spatial variation (Chica-Olmo, et al., 2004).

Once the polygon-based image representation on the lowest level of a pyramid is produced, we iteratively group polygons, sharing their contour fragments, on level *l* into larger polygonal chunks, producing level *(l+1)* of the image pyramid. Polygon agglomeration is based on Boruvka's algorithm to construct MST. Boruvka's algorithm proceeds in a sequence of stages, and in each stage it identifies a forest *F* consisting of the minimum-weight edge incident to each vertex in the graph *G*, then forms the graph $G_1 = G \backslash F$ as the input to the next stage. $G \backslash F$ denotes the graph derived from *G* by contracting edges in *F*. Boruvka's algorithm takes $O(E \log V)$ time, where *E* is number of edges and *V* is number of vertices.

The overall quality of segmentation depends on the pairwise polygon adjacency matrix, containing E_l. The attributes of edges are defined using two features: color similarity, ΔC_{ij}, and strength of the contour segment separating polygons, P_w. We evaluate the affinity w_{ij} between two neighboring polygons *i* and *j*:

$$w_{ij} = \begin{cases} \Delta C_{ij} \cdot k \cdot \exp\left(P_w / \sigma \right), \; if \; nmLODs \le N_0 \\ \left(\Delta C_{ij} + \Delta \gamma_{ij} \right) \cdot k \cdot \exp\left(P_w / \sigma \right), otherwise \end{cases}, \tag{1}$$

$$P_w = \sum_{k=1}^{N} \frac{s_k}{S} \cdot m_k \tag{2}$$

where N_0 is the number of LODs that should be constructed before variation term is taken into account, contour segment shared by two neighboring polygons consists of *N* edge fragments (Fig. 4), *S* is the length of the shared contour segment, s_k is the length of the shared edge fragment belonging to a given contour fragment, *k* and σ control the scale of polygons similarity, and m_k is the magnitude of the shared edge segment. m_k is 0 for triangle edge and non-zero for spectrally detected edge (such as Canny edges). Thus the cost of merging two polygons separated only by triangle edges is less than the cost of merging polygons separated by spectrally detected edges.
The algorithm constructs level *(l+1)* of a pyramid containing coarser image partitioning by running one Boruvka's iteration on level *l* using the evaluated color, variogram, and contour relations between the polygons. Once a coarser level is constructed, the color characterization of agglomerated polygons on level *(l+1)* is evaluated:

$$C_i^{l+1} = \sum_{k=1}^{M} \frac{A_k^l}{A_i^{l+1}} \cdot C_k^l \,. \tag{3}$$

where *M* is the number of polygons merged into polygon *i* on level *(l+1)*, C_i^l is the color of a polygon *i* on level *l*, A_i^l is the area of a polygon *i* on level *l*.

The spatial layout of the polygons changes after each agglomeration iteration. As a result, the adjacency matrix corresponding to level *(l+1)* is re-evaluated. This generation of coarser level based on finer level of a pyramid iteratively goes (Fig. 4) until dissimilarity threshold is exceeded.

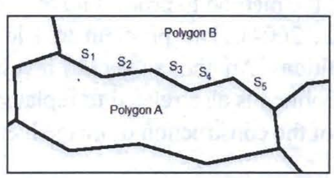

Figure 4. An example of the contour fragment shared by two neighboring polygons *A* and *B*. The shared contour fragment consists of five edge fragments of corresponding lengths s_1 through s_5.

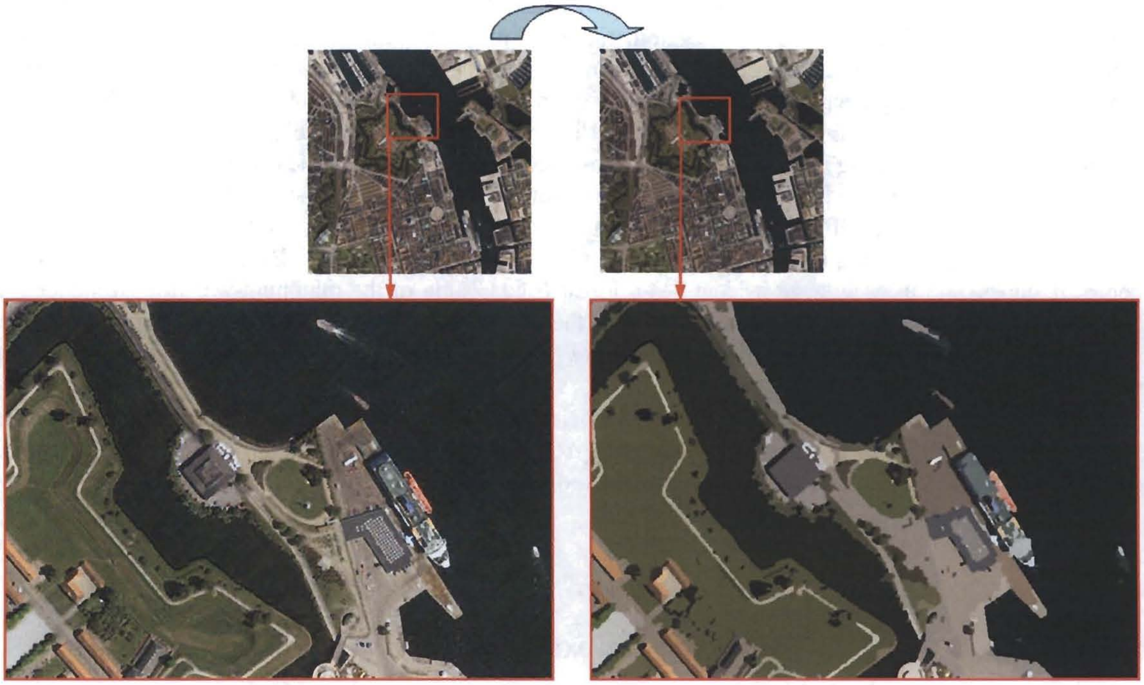

Figure 5. An example of the hierarchical image segmentation. Source of the original image: DigitalGlobe.com Top row: Original image size of size 2,512×2,525 pixels is on the left; the result of polygon agglomeration on LOD=10 is on the right, number of polygons on LOD=10 is 22,773. Bottom row: region of the original image (on the left), polygon-based representation of the same region is on the right.

Experimental Results

The presented framework was preliminary evaluated using the Berkeley Segmentation dataset* and Digital Globe satellite imagery. Figures 5 through 7 show some results of satellite image segmentation using our framework. Figs. 5 and 6 show the results of segmentation on the coarsest LOD; Fig. 7 show the result of man-made object detection based on the coarsest LOD produced by the presented framework.

Fine level of detail of a pyramid is constructed using Canny edge detector with the σ_{Canny} = 1., hysteresis low threshold = 2.5, and hysteresis high threshold = 5. We use the Triangle code** to generate triangular tessellation over the detected edge map. Color images were processed using either one of CIELab, CIELuv, or HSV spaces. N_0 is set to 2 to create two coarser LODs on top of the fine LOD before variogram is taken into account. The results are shown for CIELab space. Global threshold was set to 20 because the perceptually significant difference in color space is estimated in the range [15, 30]. Scale parameters k and σ for adjacency relations were both set to 1. On average the produced hierarchies contained 6-10 levels of detail.

CONCLUSIONS

We have presented a framework to construct a fine-to-coarse hierarchy of irregular image partitions. Experimental results support the validity of the proposed method. It uses spectral and contour relations between polygons as criteria of their agglomeration. Computational complexity of the algorithm makes it possible to use it for processing large images. Additional research is required to generalize the method to process texture, e.g. using Gabor filters, wavelets (Jain, et al., 1991; Livens, et al., 1997; Newsam, et al., 2004). The problem to address is more efficient evaluation of texture properties to attribute to the generated partitions. Another subject to investigate is an automated evaluation of optimal LODs to perform structural analysis. This problem is also related to replacing fixed scheme of selection of N_0 threshold with a scheme that depends on the progress of the construction of image hierarchy.

---

* http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/

** http://www.cs.cmu.edu/~quake/triangle.html

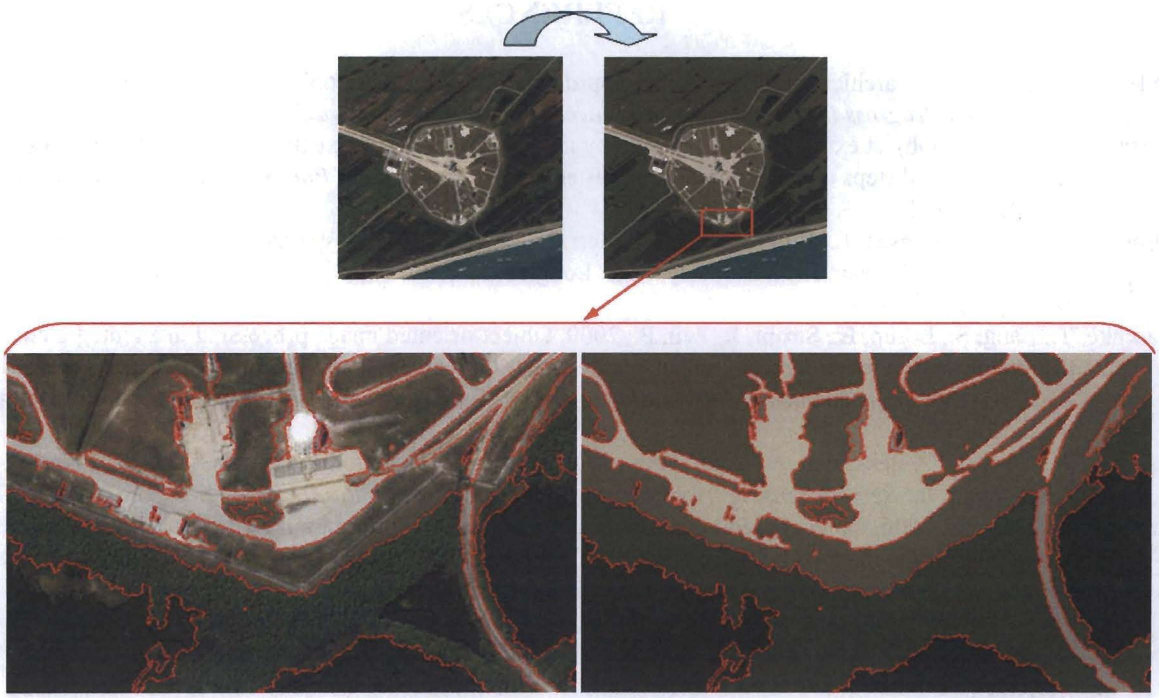

Figure 6. An example of the hierarchical image segmentation. Source of the original image: DigitalGlobe.com Top row: Original image size of size 2,897×2,891 pixels is on the left; the result of polygon agglomeration on LOD=10 is on the right, number of polygons on LOD=10 is 1,026. Bottom row: Contours of the polygons are superimposed on top of the original image (on the left), polygons attributed with color are shown on the right.

Figure 7. An example of man-made object detection using partitions on single LODs of the built hierarchy. Left: Image of nuclear facility at Arak, Iran. Source: DigitalGlobe.com. Right: the result of detection (in white). In order to achieve better detection it is necessary to perform the detection across multiple LODs and then integrate results.

REFERENCES

Anders, K.H., 2003. A hierarchical graph-clustering approach to find groups of objects, In: *Proceedings of the 5th Workshop on Progress in Automated Map Generalization.*

Baltsavias, E.P., 2004. Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems, *ISPRS Journal of Photogrammetry & Remote Sensing*, 58(3-4), 129-151.

Barnsley, M.J., Moller-Jensen, L., Barr, S.L., 2000. Inferring urban land use by spatial and structural pattern recognition, In: Donnay, J.-P., Barnsley, M.J., Longley, P.A. (Eds.) *Remote sensing and image analysis*, Routledge, pp. 115-144.

Blaschke, T., Lang, S., Lorup, E., Strobl, J., Zeil, P., 2000. Object-oriented image processing in an integrated GIS/remote sensing environment and perspectives for environmental applications. In: Cremers, A. and Greve, K. (Eds.): *Environmental Information for Planning, Politics and the Public.* Metropolis, Marburg, vol. 2, 555-570.

Blaschke, T., Burnett, C., Pekkarien, A., 2004. Image segmentation methods for object-based analysis and classification, In: *Remote Sensing Image Analysis: Including the Spatial Domain*, Springer, pp. 211-236.

Blaschke, T., Lang, S., 2006. Object based image analysis for automated information extraction - a synthesis. In: *Proceedings of the MAPPS/ASPRS 2006 Fall Conference, Measuring the Earth II: Latest Developments with Digital Surface Modeling and Automated Feature Extraction*, http://ispace.researchstudio.at/downloads/2006/173_full.pdf

Bock, M., Lessing, R., 2000. Remote sensing, formation of objects and determination of quality, In: *Internationales Symposium* "Informatik für den Umweltschutz" der Gesellschaft für Informatik (GI), Bonn, http://enviroinfo.isep.at/UI%20200/BockM300700.el.hsp.pdf

Borůvka, O. 1926. O jistém problému minimálnim. (about a certain minimal problem) *Práce Moravské Přírodvědecké Společnosti v Brně (Acta Societ. Scienc. Natur. Moravicae)*, 3(3), 37-58. (In Czech.)

Boyer, K.L., Sarkar, S., (Eds.), 2000. *Perceptual Organization for Artificial Vision Systems*, Kluwer Acad Publ.

Boykov, Y., Veksler, O., Zabin, R., 2001. Fast approximate energy minimization via graph cuts, *IEEE Trans. On Pattern Analysis and Machine Intelligence*, 23(11), 1222-1239.

Canny, J., 1986. A computational approach to edge detection, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6), 679-698.

Chen, J., Zhao, R., Li, Z., 2004. Voronoi-based k-order neighbour relations for spatial analysis, *ISPRS Journal of Photogrammetry and Remote Sensing*, 59, 60-72.

Chica-Olmo, M., Abarca-Hernandez, F., 2004. Variogram derived image texture for classifying remotely sensed images, In: *Remote Sensing Image Analysis: Including the Spatial Domain*, Springer, pp. 93-111.

Corcoran, P., Winstanley, A.C., 2007. Segmentation evaluation for object based remotely sensed image analysis, In: *GISRUK 2007 Geographical Information Science Research UK Conference*, National Centre for Geocomputation, NUI Maynooth, Ireland.

Felzenszwalb, P.F., Huttenlocher, D.P., 1998. Image segmentation using local variation, *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition*, 98-104.

Fowlkes, C., Martin, D., Malik, J., 2003. Learning affinity functions for image segmentation: combining patch-based and gradient-based approaches, In: *Proceedings of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2, 54-61.

Fuller, R.M., Smith, G.M., Thomson, A.G., 2004. Contextual analyses of remotely sensed images for the operational classification of land cover in United Kingdom, In: S.M., de Jong, F.D. van der Meer (Eds.) *Remote Sensing Image Analysis: Including the Spatial Domain*, Springer, 271-290.

Haxhimusa, Y., Kropatsch, W. G., 2004. Segmentation graph hierarchies, *Lecture Notes in Computer Science*, 3138, 343-351.

Haxhimusa, Y., Ion, A., Kropatsch, W.G., 2006. Comparing hierarchies of segmentations: Humans, normalized cut, and minimum spanning tree, In: *Digital Imaging and Pattern Recognition*, 95-103.

Herold, M., Gardner, M.E., Roberts, D.A., 2003a. Spectral resolution requirements for mapping urban areas, *IEEE Trans. on Geoscience and Remote Sensing*, 41(9), 1907-1919.

Herold, M., Goldstein, N., Clarke, K.C., 2003b. The spatiotemporal form of urban growth: measurement, analysis and modeling, *Remote Sensing of Environment*, 86(3), 286-302.

Herold, M., Roberts, D. A., Gardner, M. E., Dennison, P. E., 2004. Spectrometry for urban area remote sensing – Development and analysis of a spectral library from 350 to 2400 nm, *Remote Sensing of Environment*, 91(3-4), 304-319.

Hoffmann, A., van der Vegt, J., 2001. New sensor systems and new classification methods: Laser- and digital camera- data meet object-oriented strategies, *Geo-Information Systems*, 6/01, S. 18-23.

Horowitz, S.L., Pavlidis, T., 1976. Picture segmentation by a tree traversal algorithm, *Journal of the Association for Computing Machinery*, 23(2), 368-388.

Jain, A.K., Farrokhnia, F., 1991. Unsupervised texture segmentation using Gabor filters, *Pattern Recognition*, 24(12), 1167-1186.

Jaromczyk, J.W., Touyssaint, G.T., 1992. Relative neighborhood graphs and their relatives, *Proceedings of the IEEE*, 80(9), 1502-1517.

Jolion, J.-M., Montanvert, A., 1992. The adaptive pyramid, a framework for 2D image analysis, *CVGIP: Image Understanding*, 55(3), 339-348.

Kropatsch, W.G., Haxhimusa, Y., Ion, A., 2007. Multiresolution image segmentation in graph pyramids. In: *Applied Graph Theory in Computer Vision and Pattern Recognition,* Kandel, A., Bunke, H. H., Last, M. (Eds.) 52, 3-42.

Kruskal, J.B.J., 1956. On the shortest spanning subtree of a graph and the traveling salesman problem, In: *Proc. Am. Math. Soc.*, 7, 48-50.

Kunz, D., Schilling, K.-J., Vogtle, T., 1997. A new approach for satellite image analysis by means of a semantic network, In: Förstner, W., Plumer, L. (Eds.), *Semantic modeling for the acquisition of topographic information from images and maps: SMATI 97*, Springer, 20-36.

Livens, S., Scheunders, P., de Wouwer, G.V., Dyck, D.V., 1997. Wavelets for texture analysis, In: *Proceedings of the 6th Int. Conference on Image Processing and Its Applications*, vol. 2, pp. 581-584.

Liu, Y., Guo, Q., Kelly M., 2008. A framework of region-based spatial relations for non-overlapping features and its application in object based image analysis, *ISPRS Journal of Photogrammetry & Remote Sensing*, 63, 461-475.

Lowell, K., Gold, C., Edwards, G., 1992. The next generation of digital spatial technologies: error utilization, thematic map and remote sensing integration, spatial operators, and data structures, In: *Proceedings GIS'92 Symposium*.

Mustière, S., Moulin, B., 2002. What is spatial context in cartographic generalization?, In: *Proceedings of Symposium on Geospatial Theory, Processing and Applications*, Ottawa, Canada.

Montanvert, A., Meer, P., Rosenfeld, A., 1991. Hierarchical image analysis using irregular tessellations, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(4), 307-316.

Neubert, M., Herold, H., Meinel, G., 2004. A comparison of segmentation programs for high resolution remote sensing data, In: *Proceedings of the XXth ISPRS Congress, Int. Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, Vol. XXXV-B4, pp. 1097-1102, http://www.cartesia.org/geodoc/isprs2004/comm4/papers/506.pdf

Neubert, M., Herold, H., Meinel, G., 2006. Evaluation of remote sensing image segmentation quality - further results and concepts, In: *Proceedings of the 1st Int. Conference on Object-based Image Analysis (OBIA 2006)*, http://www.commission4.isprs.org/obia06/papers.htm

Neubert, M., Herold, H., Meinel, G., 2008a. Assessing image segmentation quality - Concepts, methods and application, In: Blaschke, T.; Hay, G., Lang, S. (Eds.): *Object-Based Image Analysis - Spatial concepts for knowledge-driven remote sensing applications, Lecture Notes in Geoinformation and Cartography*, Springer, Berlin, 769-784.

Neubert, M., Herold, H., Meinel, G., 2008b. Assessment of remote sensing image segmentation quality, In: *Proceedings of the Int. Conference on Geographic Object Based Image Analysis (GEOBIA 2008)*, http://homepages.ucalgary.ca/~gjhay/geobia/Proceedings/Sessions/GEOBIA_Proceedings_RP10_FINALc.pdf

Newsam, S., Wang, L., Bhagavathy, S., Manjunath, S., 2004. Using texture to analyze and manage large collections of remote sensed image and video data, *Applied Optics*, 43(2), 210-217.

Prasad, L., Skourikhine, A.N., 2006a. Vectorized image segmentation via trixel agglomeration, *Pattern Recognition*, 39(4), 501-514.

Prasad, L., Skourikhine, A.N., 2006b. Vectorized image segmentation via trixel agglomeration, *U.S. Patent No.* 7127104.

Prim, R.C., 1957. Shortest connection networks and some generalizations, *The Bell System Technical Journal*, 36, 1389-1401.

Regnauld, N., 2001. Contextual building typification in automated map generalization, *Algorithmica*, 30, 312-333.

Schiewe, J., Tufte, L., Ehlers, M., 2001. Potential and problems of multi-scale segmentation methods in remote sensing, *Geo-Information Systems*, 14(6), 34– 39.

Sharma, O., Mioc, D., Anton, F., 2008. Polygon feature extraction from satellite imagery based on colour image segmentation and medial axis, In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Siences*, Vol. XXXVII, Part 3a, 235-240.

Sharon, E., Galun, M., Sharon, D., Basri, R., Brandt, A., 2006. Hierarchy and adaptivity in segmenting visual scenes, *Nature*, 442, 810-813.

Shewchuk, J.R., 1996. Triangle: engineering a 2D quality mesh generator and Delaunay triangulator, *Lecture Notes in Computer Science*, 1148, 203-222.

Shi, J., Malik, J., 2000. Normalized cuts and image segmentation, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(8), 888-905.

Skourikhine, A.N., 2006. Proximity graph analysis for linear networks extraction from high-resolution satellite imagery, In: *Signal Processing, Sensor Fusion, and Target Recognition XV Conference, Proc. SPIE*, vol. 6235.

Skurikhin, A.N., 2008. Proximity graphs based multi-scale image segmentation. *Proceedings of the 4th Int. Symposium on Visual Computing, Lecture Notes in Computer Science*, 5358, 298-307.

Skurikhin, A.N., Volegov, P.L., 2008. Object-oriented hierarchical image vectorization, In: *Proceedings of the Int. Conference on Geographic Object Based Image Analysis (GEOBIA 2008)*, http://homepages.ucalgary.ca/~gjhay/geobia/Proceedings/Sessions/GEOBIA_Proceedings_RP10_FINALc.pdf

Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A., Jain, R., 2000. Content-based image retrieval at the end of the early years, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22(12), 1349-1380.

Strat, T.M., Fischler, M.A., 1991. Context-based vision: Recognizing objects using information from both 2-D and 3-D imagery, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(10), 1050-1065.

Taubenböck, H., Esch, T., Roth, A., 2006. An urban classification approach based on an object-oriented analysis of high resolution satellite imagery for a spatial structuring within urban areas. In: *Proceedings of the 1st EARSel Workshop of the SIG Urban Remote Sensing.*

Torralba, A., 2003. Contextual priming for object detection, *Int. Journal of Computer Vision*, 53(2), 169-191.

Toussaint, G.T., 1980. The relative neighborhood graph of a finite planar set, *Pattern Recognition*, 12, 261-268.

Quint, F., Sties, M., 1995. Map-based semantic modeling for the extraction of objects from aerial images, In: Grün, A., (Ed.), *Automatic extraction of man-made objects from aerial and space* images, 307-316.

Unnikrishnan, R., Pantofaru, C., Hebert., M., 2007. Toward Objective Evaluation of Image Segmentation Algorithms, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 29(6), 929-944.

Walter, V., 2004. Object-based classification of remote sensing data for change detection, *ISPRS Journal of Photogrammetry & Remote Sensing*, 58, 225-238.

Weiss, Y., 1999. Segmentation using eigenvectors: a unifying view, In: *Proceedings of the 7th IEEE Int. Conf. on Computer Vision*, 2, 975-982.

Wertheimer, M., 1958. Principles of perceptual organization, In: *Readings in Perception*, Beardslee, D., Wertheimer, M., (Eds.), D. Van Nostrand, Princeton, NJ, 115-135.

Wu, Z., Leahy, R., 1993. An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 15(11), 1101-1113.

Xu, Y., Uberbacher, E. C., 1997. 2D image segmentation using minimum spanning trees, *Image and Vision Computing*, 15, 47-57.

Zahn, C. T., 1971. Graph-theoretic methods for detecting and describing gestalt clusters, *IEEE Transactions on Computers*, 20(1), 68-86.

Zhan, Q., Molenaar, M., Tempfli, K., 2002. Finding spatial units for land use classification based on hierarchical image objects, *Proceedings of the Symposium on Geospatial Theory, Processing and Applications*, In: *The International Archives of Photogrammetry and Remote Sensing*, vol. 34, Part 4, 263-268.

Zhang, Y., 1996. A survey on evaluation methods for image segmentation. *Pattern Recognition,* 29(8), 1335-1346.