LA-UR- 09-01410

Approved for public release; distribution is unlimited.

Title: Man-made objects cuing in satellite imagery

Author(s): Skurikhin, A.N.

Intended for: 2009 SPIE Defense, Security and Sensing: Signal Processing, Sensor Fusion, and Target Recognition XVIII Conference; Orlando, Florida, April 2009

Los Alamos NATIONAL LABORATORY EST. 1943

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the Los Alamos National Security, LLC for the National Nuclear Security Administration of the U.S. Department of Energy under contract DE-AC52-06NA25396. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

Form 836 (7/06)

Man-made objects cuing in satellite imagery

Alexei N. Skurikhin*
MS D436, Space and Remote Sensing Group, Los Alamos National Laboratory, Los Alamos, NM, USA 87545

ABSTRACT

We present a multi-scale framework for man-made structures cuing in satellite image regions. The approach is based on a hierarchical image segmentation followed by structural analysis. A hierarchical segmentation produces an image pyramid that contains a stack of irregular image partitions, represented as polygonized pixel patches, of successively reduced levels of detail (LODs). We are jumping off from the over-segmented image represented by polygons attributed with spectral and texture information. The image is represented as a proximity graph with vertices corresponding to the polygons and edges reflecting polygon relations. This is followed by the iterative graph contraction based on Boruvka's Minimum Spanning Tree (MST) construction algorithm. The graph contractions merge the patches based on their pairwise spectral and texture differences. Concurrently with the construction of the irregular image pyramid, structural analysis is done on the agglomerated patches. Man-made object cuing is based on the analysis of shape properties of the constructed patches and their spatial relations. The presented framework can be used as pre-scanning tool for wide area monitoring to quickly guide the further analysis to regions of interest.

Keywords: irregular pyramid, segmentation, minimum spanning tree, Boruvka, Delaunay triangulation, man-made, cuing, level-of-detail

1. INTRODUCTION

With very high-resolution satellite and aerial surveillance becoming ubiquitous and data is being generated at an enormous rates, automated cuing of analyst attention on most interesting regions of the original image has become a subject of intensive research. Object-oriented image segmentation, detection and classification of regions containing man-made objects, such as road networks and buildings, have been given a considerable attention (for a comprehensive review see Ridd, M., K., et al., 2006). However, the complete automation of this process is still an unsolved problem. While many algorithms and approaches have been developed, the manual intervention of the analyst in validating segmented objects is needed. State of the art on methods of road network extraction and reconstruction is presented in (Mena, 2003). The surveyed methods are broadly classified according to the preset objective, the applied extraction technique, and the type of used sensor. The paper (Mayer, 1999) surveys the state of the art on building extraction from aerial imagery. Assessment of different approaches is performed according to complexity of the employed models and strategies. The model comprises the function of objects, material properties, 2D and 3D geometry, scales and levels of abstraction, structures of parts, and local and global context. The strategy consist of grouping based on geometric and topological regularities, focusing on different scales, context driven search and segmentation, generation of evidence from structures of parts, and fusion of data and algorithms. The paper (Baltsavias et al., 2004) describes approaches to extract cartographic objects and to update GIS. The analysis focuses on extraction of buildings and roads, and reviews recent knowledge-based image analysis trends, knowledge representation and augmenting. A review of techniques for linear feature detection in images is presented in (Quackenbush, 2004). Reviewed techniques include mathematical morphology, Hough transform, multi-resolution edge detection, template matching, dynamic programming for edge linking, and rule-based classification. Most of these techniques were developed for extracting roads.

Very high resolution (VHR) satellite imagery, such as QuickBird and WorldView satellite images of 0.61m and 0.5m spatial resolution, respectively, reveals many details previously unobservable in satellite images; however, at the same time, this type of data presents a real challenge. Increases in spatial resolution change object appearance and 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

*alexei@lanl.gov; phone 1 505 667-5067; fax 1 505 664-0362

(Baltsavias, 2004; Schiewe, et al., 2001). The challenge is due to the fact that the elements of interest are not individual pixels, but pixel patches (a.k.a. superpixels, polygons). 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). Therefore, the use of structural features, such as shape, becomes very important for differentiation of spectrally similar objects. Objects in satellite imagery are often networks of spatially distributed entities of varying sizes (e.g. an industrial site) and can be represented in more than one way, and using more than one level of details (LODs). Thus, hierarchic object-oriented segmentation (Blaschke, et al., 2006; Hoffmann, et al., 2001; Liu et al., 2008; Walter, 2004; Zhan, et al., 2002) and spatial analysis (Barnsley, et al., 2000; Blaschke, et al., 2000, 2004; Bock, et al., 2000; Fuller, et al., 2004; Strat, et al., 1991; Lowell et al. 1992; Mustière et al., 2002; Herold et.al, 2003a) are indispensable for a robust satellite image interpretation, map updating, cartographic generalization, and change analysis.

This paper presents a framework for man-made object cuing based on hierarchical image segmentation and feature extraction. The framework constructs an irregular pyramid that contains a stack of polygonized images of successively reduced LODs. The polygonized images are represented as sets of polygonized pixel patches (polygons) attributed with spectral, textural, and structural characteristics. The pyramid is built upon an oversegmented image. Oversegmentation is achieved by the use of the constrained Delaunay triangulation of edges detected in the original image, followed by filtering of generated triangle edges. This closes gaps between edge fragments and creates closed contours outlining an initial set of seed polygons. A polygonized image is then represented as a proximity graph with vertices corresponding to the polygons and edges reflecting polygon relations. This is followed by the iterative graph contraction using Boruvka's MST construction algorithm that merges the patches until dissimilarity criteria are exceeded. Concurrently with the construction of the image pyramid, structural analysis is done on the agglomerated patches. Man-made object cuing is based on the analysis of shape properties of the constructed patches at multiple LODs. Once cuing to regions containing man-made objects is done, the further more detailed analysis can be quickly guided to such regions of interest.

In the next section we present related research in the areas of man-made object detection and hierarchical image segmentation using proximity graphs. Section 3 explains our framework for image segmentation, and feature extraction. In this section we also present experimental results using QuickBird satellite imagery. Finally, conclusions and future work are given in section 4.

2. RELATED WORK

2.1 Man-made object detection

Most of the work in VHR imagery centers on the problem of recognizing specific objects, such as roads and buildings, using spectral and texture features. Less work has been published about the robust approaches to detect and characterize regions containing generic class of man-made objects using structural features.

A recent work of (Inglada, 2007) describes a supervised learning method to detect generic classes of man-made objects in satellite images with 2.5 m resolution. Detection is defined as finding a small region containing an object. High number of geometric descriptors are extracted to cope with the diversity of possible object appearance. This is followed by SVM-based classification that also deals with high dimensionality of the feature vector. To build a system which is robust to spectral variation, only geometric features were used. The selected classes of man-made objects are well defined by spatial relations of their edges. This makes possible to avoid spectral characteristics. The considered classes included isolated buildings, paths and tracks, crossroads, bridges, wide roads, highways, round-abouts, narrow roads, railways, and surburbs. The extracted feature vector includes Hu invariants (Hu, 1962), coefficient of the Fourier-Mellin transform, and high-level geometry features related to analyzed pixel patch. The high level features comprises the entropy of the orientations of the alignments and the histograms of the distances between selected image elements and their sizes. Since the sequential scanning of entire image can be time consuming, it is necessary to use a focusing strategy cuing an analysis on interesting regions. Further research is also required to extract features more suitable for VHR satellite images.

(Mueller et al., 2004) presents an object-oriented image segmentation approach with focus on shape analysis. The approach is based on the combination of edge- and region-based segmentation techniques. Shape features, such as straight edges, are extracted and control region agglomeration process. Initially, edges are extracted at multiple scales and, then they are evaluated using a proposed measure of straightness. Edges meeting prespecified criteria on edge contrast, straightness, length are kept. Further analysis on the remaining edges preserves edges lying each from other within a tolerable proximity; this results in further edge set reduction. The reduced edge set complements region agglomeration that is based on the use of mean intensities of growing regions. This is done by introducing two criteria that should be satisfied for regions to be merged. The first one is about whether the pixels are on opposite sides of the edge, and the second one prevents regions merging through a bottleneck, when regions are connected through a small gap between preserved edges. The approach focuses on the extraction of agricultural fields; thus it assumes that the objects are large and rather homogeneous areas with respect to gray-scale intensities. The approach is tested on images from different satellites with spatial resolutions 1m, 4.5m, and 5.8 m.

In (Molinier, 2007) a content-based image retrieval system for the analysis of remote sensing imagery is described. Detection of regions containing man-made objects is based on a combination of self-organizing maps (Kohonen, 2001) and is validated on multispectral (2.4m resolution) and panchromatic (0.6m resolution) QuickBird imagery. Prior to analysis, each satellite image is split into small regions (imagelets), which are attributed with feature vectors. The feature vector comprises average red, green, blue, color moments, texture (extracted from panchromatic image), *xy*-coordinates, NDVI histogram, and edges histogram (extracted from panchromatic image). While experimental evaluation of the approach showed its efficiency on the used data, further research is necessary to investigate the influence of the imagelet size, sampling of the original image with the imagelets, and spatial relations between neighboring imagelets.

(Iqbal et al, 2002) proposes computer vision framework to exploit image structure for retrieval of images containing man-made structures from a database. The framework extracts features which are evidence of the presence of man-made objects. This is followed by classification of these features to detect images with man-made objects. The used features are straight line segments, a set of lines terminating at a common point (co-terminations), "L" and "U" junctions, parallel lines, and polygons. Grouping of line segments to extract appropriate features is based on perceptual grouping rules known from visual psychology (Wertheimer, 1958). Polygons are reconstructed using search for the fundamental circuits of a co-termination graph. Each fundamental circuit represents a polygon, where edges on this circuit correspond to line segments. Each polygon is validated against a number of pre-specified criteria. The approach is validated using outdoor imagery.

One of the applications that exploits man made object detection and characterization is nonproliferation verification using satellite imagery. Overview of such application of VHR commercial satellite imagery is presented in (Pabian, 2008). This book chapter describes approaches for the interpretation of commercial satellite imagery, provides a very extensive reference list and links to a variety of publicly available resources, and discusses practicalities and limitations of such imagery for detecting undeclared nuclear activities. (Jasani et al, 2002) also examines the role of satellite imagery in preventing nuclear proliferation, and discusses the use of satellites by the International Atomic Energy Agency (IAEA). The book contains a collection of contribution from a number of world-renowned experts in a variety remote sensing fields. A complete object-based image analysis framework that uses man-made object detection for treaty verification is proposed and investigated in (Nissbaum et al, 2008). The framework is based on multi-scale image segmentation that produces an image hierarchy of object-oriented pixel patches, which is analyzed by semantic modeling to detect and characterize man-made structures. While the presented framework holds a promise for the considered application, it's efficiency is severely limited by the quality of image segmentation. The authors write in their conclusions the following (on page 144): "Since, however, the starting point for all results produced by object-based image analysis is the segmentation, the latter decisively influences the possibilities and limitations. At the moment the automated segmentation algorithm used ... limits the general potential of the image analysis since the segmented object boundaries do not exactly correspond to the real world objects."

The above studies corroborate difficulty of achieving good quality object-oriented segmentation, especially when more accurate object segmentation is required for semantic image interpretation. This is a well known outstanding problem in computer vision. Structural and high-level analysis that follows segmentation relies on the quality of extracted pixel patches due to the use of shape characteristics and spatial relations of pixel patches. Therefore, the quality of segmenting (or grouping) of image elements, such as pixels, into larger structures directly affects the performance of object detection and characterization.

2.2 Proximity graphs based image segmentation

Delaunay triangulation (DT) 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) specify the formation of perceptually significant and visually attractive patterns. 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. Different inter-element relational attributes reflecting these criteria are usually condensed into composite weights, "generalized costs", to make a problem of groping elements into objects computationally tractable. Further example of spatial analysis includes land use classification in (Zhan et al, 2002) that is based on Delaunay triangulation that derives spatial relations between the detected image objects. (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. (Sharma et al, 2008) uses the Delaunay and the Voronoi graphs to extract features characterizing road and hydrographic networks.

While spatial analysis techniques based on proximity graphs are widely used and advanced, they are based on the assumption that objects of interest are available, either as a result of supervised spectral and texture-based classification of original imagery or unsupervised segmentation. Bottom-up unsupervised image segmentation approach holds a lot of promise for image interpretation, at the same time, it is an extremely difficult problem when the environment is unconstrained and automated segmentation is required. The principle of this approach is that each initial pixel patch will grow until no more similar patches can be added to it. We summarize here few techniques that create an initial image representation consisting of pixels patches (superpixels) that can serve as starting data set for the bootm-up agglomeration process. (Prasad et al, 2006a, 2006b) propose a computationally efficient image segmentation framework that uses the constrained DT of the image edge set followed by selective filtering of generated triangulation edges. This creates an oversegmented image. It contains polygonized pixel patches outlined by closed contours consisting of detected edges and preserved triangulation edges (which close gaps between detected edges). A related example of how DT can be used to complete piecewise linear contour approximations is given in (Ren at al, 2005) . It should be noted that there are other than triangulation-based approaches for image pre-segmentation and creation of superpixels. They can be created using such approaches as watershed based segmentation (Vincent et al, 1991), and Mean Shift that is a kernel based density estimation technique (Fukunaga et al, 1975; Cheng, 1995; Comaniciu et al, 2002). Examples of the approaches of how to build a segmentation from the originally oversegmented image include (Li et al, 2004; Chefd'hotel et al, 2007; Stawiaski et al, 2008; Hanbury, 2008).

We emphasize a category of algorithms that seek optimal image hierarchic partitioning through a sequence of local computations based on proximity graphs, specifically MST. 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. (Morris et al, 1986) proposes hierarchic image segmentation approach based on Kruskal's MST construction algorithm (Kruskal, 1956), starting from a regular pixel grid. Hierarchic merging of pixel patches is controlled by updated intensity dissimilarities between the agglomerated patches. (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 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 merged with the larger

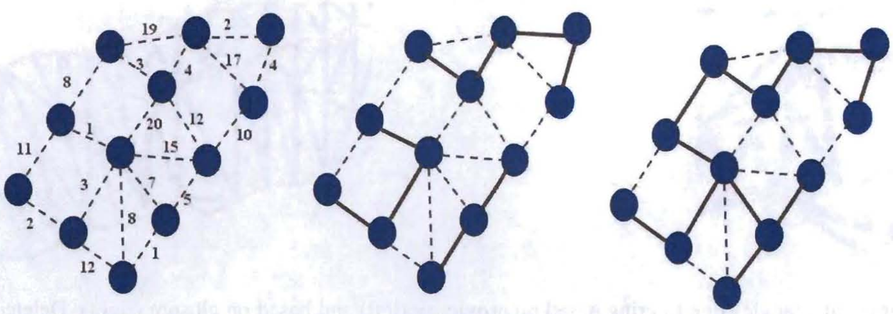

Fig. 1. Iterations of Boruvka's MST reconstruction (from left to right). MST edges are shown in solid black lines; numbers represent similarities between the nodes, the lower the number, the more similar the nodes are.

neighbors. 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; Nešetřil et al, 2001) 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.

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 the proximity graphs can now be applied to extract and exploit contextual relations between the objects across multiple LODs. The proposed framework, as well as the above proximity graph based approaches, can also be generalized to process pixel patches (superpixels) produced by alternative non-triangulation based techniques.

3. OBJECT ORIENTED IMAGE SEGMENTATION AND ANALYSIS

In a polygon-based image pyramid, each LOD 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.

3.1 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). This is followed by constrained Delaunay triangulation (CDT) (Schewchuk, 1996) where the detected edges are used as constraints for the triangulation. Thus, the CDT tessellation grid is adapted to the image content, since triangle vertices and edges reflect the structure and spatial

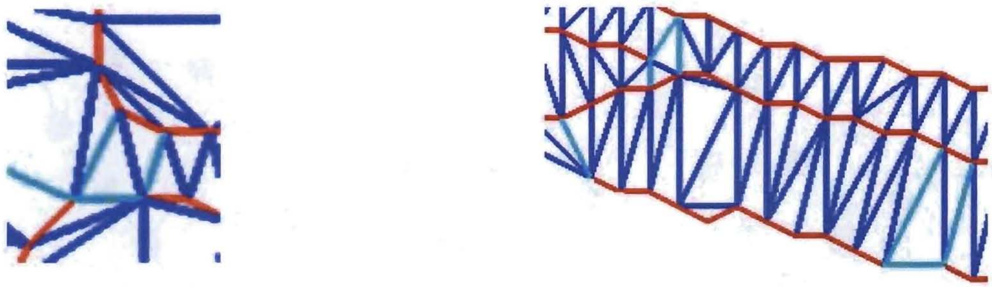

Fig. 2. Examples of triangle edge filtering based on proximity (left) and based on closure (right). Deleted triangle edges are shown in blue, and the kept triangle edges (closing the gaps between Canny edges) are shown in cyan color.

adjacency of the detected edges. CDT generated triangular mesh is then processed by edge filtering. The filtering keeps constraints (the detected edges) and selectively deletes generated triangle edges. Triangle edge filtering uses the principles of visual perception, such as proximity and closure. Proximity filters out triangle edges based on their length (Fig. 2, left). As a result, the detected edges that are spatially close to each other are linked by the kept shortest triangle edge. 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 (Fig. 2, right). This results in a set of closed contours consisting of combination of the generated triangle edges and spectrally detected edges. Finally, a graph traversal algorithm groups triangles within the constructed closed contours into polygons. 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 polygons. The drastic reduction in the amount of data, number of pixels to number of generated polygons, by 20-80 times provides significant gains in computational efficiency for further analysis.

3.2 Hierarchical image segmentation

Building larger polygons on top of the produced seed polygons has the following 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. We extend our previous work (Skurikhin, 2008; Skurikhin et al, 2008) by taking into account variogram-based estimation of 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 (Fig. 3). 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.

The overall quality of segmentation depends on the pairwise polygon adjacency matrix, containing E_l. The attributes of edges are defined using three features: color similarity, ΔC_{ij}, variogram difference, $\Delta\gamma_{ij}$, 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 variogram 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

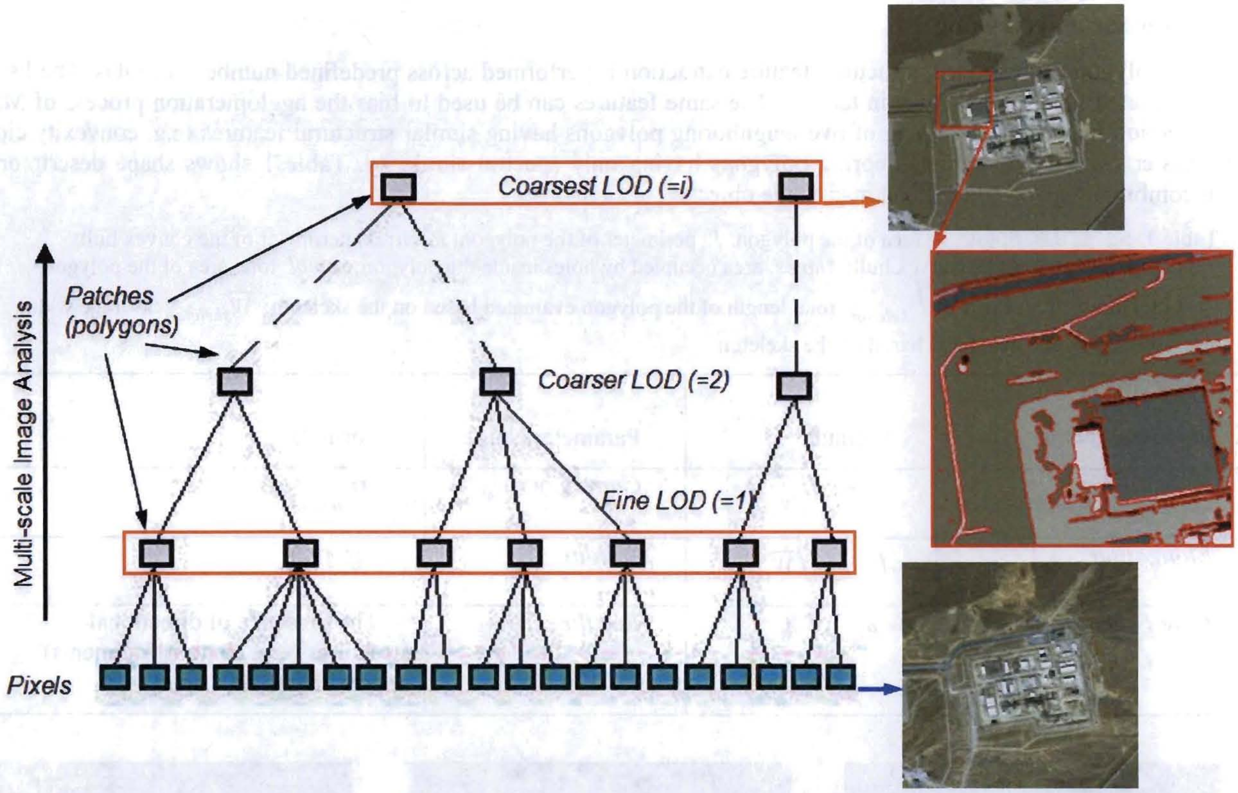

Fig. 3. Hierarchical image segmentation the produces an irregular pyramid of fine-to-coarse LODs. An example of segmentation is shown on the right. Contours of the constructed patches are shown in red.

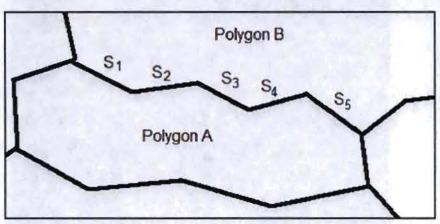

Fig. 4. An example of the contour fragment shared by two neighboring polygons *A* and *B*. The shared contour fragment consists of five edge fragments of corresponding lengths s_1 through s_5.

spectrally detected edge. 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 until pre-specified dissimilarity threshold is exceeded.

3.3 Man-made objects cuing

Once the polygons constructed, structural feature extraction is performed across predefined number of LODs. The list of the investigated features is shown in table 1. The same features can be used to bias the agglomeration process of MST reconstruction by making grouping of two neighboring polygons having similar structural features (e.g. convexity close to 1.) easier contrary to the neighboring polygons having only spectral similarity. Table 1 shows shape descriptors , which combination is used to detect man-made objects.

Table 1. Shape descriptors. *A*: area of the polygon; *P*: perimeter of the polygon; *Pconvex*: perimeter of the convex hull; *Aconvex*: area of the convex hull; *Aholes*: area occupied by holes inside the polygon; *Atotal*: total area of the polygon including *Aholes* and *A*; $l_{skeleton}$: total length of the polygon evaluated based on the skeleton; $\overline{w}_{skeleton}$:average width of the polygon evaluated based on the skeleton.

Parameter Name	Formula	Parameter Name	Formula
Form factor	$4 \cdot \pi \cdot A / P^2$	*Convexity*	P_{convex} / P
Elongation	$l_{skeleton} / \overline{w}_{skeleton}$	*Solidity*	A / A_{convex}
Hole Fraction	A_{holes} / A_{total}	*Rectilinearity*	The presence of directional structures (e.g. contour fragments) approximately 90° apart

Fig. 5. An example of hierarchic image segmentation. Left: a satellite image of the Yongbyon nuclear reactor in North Korea, 806×762 pixels (credit: © DigitalGlobe). Center: segmented image on LOD=6, 191 polygons (outlined with red contours). Right: segmented image on LOD=8, 161 polygons. Initial fine LOD=1 (not shown) contains 30,002 polygons.

3.4 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 and man-made object cuing using our framework. The results of man-made object cuing are shown using prespecified LODs. The challenge in application of the structural analysis to the polygons lies in automated choice of LODs, on which the analysis has to be applied to. Fine level of detail of a pyramid is constructed using Canny edge detector with the σ_{Canny} = 1., hysteresis low threshold = 0., and hysteresis high threshold = 5. We use the Triangle code** to generate triangular tessellation over the

* http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/segbench/
** http://www.cs.cmu.edu/~quake/triangle.html

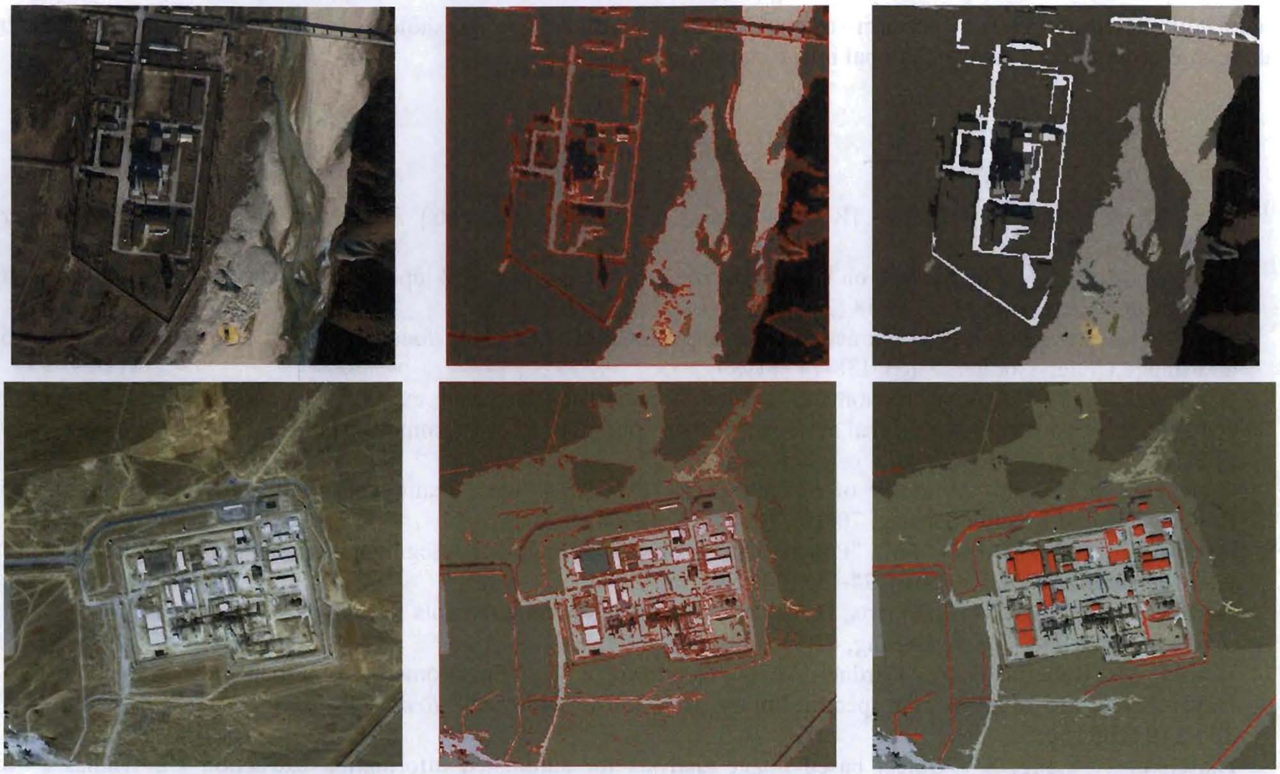

Fig. 6. An example of man-made object detection.

Top row. Left: a satellite image of the Yongbyon nuclear reactor in North Korea, 806×762 pixels (credit: DigitalGlobe). Center: segmented image on LOD=8, 161 polygons (outlined with red contours). Right: result of man-made object detection on LOD=8, detected man-made objects are shown in white color.

Bottom row. Left: a satellite image of Arak nuclear site in Iran, 1528×1487 pixels (credit: DigitalGlobe). Center: segmented image on LOD=10, 1374 polygons (outlined with red contours). Right: result of man-made object detection on LOD=10, detected man-made objects are shown in red color.

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 LODs.

Total complexity is $O(N\log N + V\log E)$ (where N is number of used edge points, V is number of initial polygons, E is number of relations between the polygons), time consumed depends on how textured the image is; the more texture in the mage, the more edges will be detected, the more time will be consumed.

4. CONCLUSIONS

We have developed and presented a segmentation framework to construct a fine-to-coarse hierarchy of irregular image partitions that are utilized by structural analysis to detect man-made structures. Experimental results support the validity of the proposed approach. It uses spectral, textural, 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. The outstanding problems are development of better texture characterization that is attributed to polygonal pixel patches, automated choice of LODs to perform structural analysis, and adaptive thresholding of the agglomeration process by taking into account both local and global image statistics.

REFERENCES

[1] Ridd, M., K., Hipple, J. D. (Eds.), [Remote Sensing of Human Settlements], American Society for Photogrammetry and Remote Sensing (2006).

[2] Mena, J., B., "State of the art on automatic road extraction for GIS update: a novel classification," Pattern Recognition Letters, 24, 3037-3058 (2003).

[3] Mayer, H., "Automatic object extraction from aerial imagery - A survey focusing on buildings," Computer Vision and Image Understanding, 74(2), 138-149 (1999).

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

[5] Quackenbush, L., J., "A review of techniques for extracting linear features from imagery," Photogrammetric Engineering and Remote Sensing, 70(12), 1383-1392 (2004).

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

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

[8] Herold, M., Roberts, D. A., Gardner, M. E., Dennison, P. E., "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 (2004).

[9] Blaschke, T., Lang, S., "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 (2006).

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

[11] Liu, Y., Guo, Q., Kelly M., "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 (2008).

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

[13] Zhan, Q., Molenaar, M., Tempfli, K., "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 (2002).

[14] 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.

[15] 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.

[16] 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, 211-236.

[17] 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

[18] 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.

[19] 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.

[20] 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*.

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

[22] Inglada, J., "Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric features," ISPRS Journal of Photogrammetry and Remote Sensing, 62, 236-248 (2007).

[23] Hu, M.,K., "Visual pattern recognition by moment invariants," IEEE Transactions on Information Theory 8 (2), 179–187 (1962).

[24] Mueller, M., Segl, K., Kaufmann, H., "Edge- and region-based segmentation technique for the extraction of large, man-made objects in high-resolution satellite imagery," Pattern Recognition, 37, 1619-1628 (2004).

[25] Molinier, M., Laaksonen, J., Hame, T., "Detecting man-made structures and changes in satellite imagery with a content-based information retrieval system built on self-organizing maps," IEEE Trans. on Geoscience and Remote Sensing, 45(4), 861-874 (2007).

[26] Kohonen, T., [Self-Organizing Maps], Berlin, Germany, Springer-Verlag, (2001).

[27] Iqbal, Q., Aggarwal, J., K., "Retrieval by classification of images containing large manmade objects using perceptual grouping," Pattern Recognition, 35, 1463-1479 (2002).

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

[29] Pabian, F., *Chapter 12. Commercial satellite imagery: Another tool in the nonproliferation verification and monitoring toolkit*, In: Doyle, J., E., (Ed.), Nuclear Safeguards, Security, and Nonproliferation, Elsevier Inc., (2008).

[30] Jasanni, B., Stein, G., 2002. Commercial satellite Imagery. A tactic in nuclear weapons deterrence, Springer.

[31] Nissbaum, S., Menz, G., Object-based Image Analysis and Treaty Verification. New Approaches in Remote Sensing - Applied to Nuclear Facilities in Iran, Springer, 2008.

[32] Toussaint, G.T., "The relative neighborhood graph of a finite planar set," *Pattern Recognition*, 12, 261-268 (1980).

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

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

[35] Regnauld, N., "Contextual building typification in automated map generalization," *Algorithmica*, 30, 312-333 (2001).

[36] Zhan, Q., Molenaar, M., Tempfli, K., "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 (2002).

[37] Skourikhine, A.N., "Proximity graph analysis for linear networks extraction from high-resolution satellite imagery," In: *Proc. SPIE*, vol. 6235 (2006).

[38] Sharma, O., Mioc, D., Anton, F., "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 (2008).

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

[40] Prasad, L., Skourikhine, A.N., "Vectorized image segmentation via trixel agglomeration," U.S. Patent No. 7127104, 2006.

[41] Ren, X., Fowlkes, C., Malik, J., "Mid-level cues improve boundary detection," Report No. UCB/CSD-5-1382, Computer Science Division, Univ. of California, Berkeley, California, March 2005.

[42] Vincent, L., Soille, P., "Watersheds in digital spaces an efficient algorithm based on immersion simulations," IEEE Trans. On Pattern Analysis and Machine Intelligence, 13(6), 583-598, 1991.

[43] Fukunaga K., Hostetler L.D., 1975. "The estimation of the gradient of a density function," *IEEE Trans Info Theory*, 21, 32–40.

[44] Cheng, Y., 1995. "Mean shift, mode seeking, and clustering," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 17, 790-799.

[45] Comaniciu, D., Meer, P., "Mean-Shift: A robust approach toward feature space analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, 24(5), 603-619 (2002).

[46] Li, Y., Sun, J., Tang, C.-K., Shum, H.-Y., 2004. "Lazy snapping", *Proc. of SIGGRAPH*, 23(3), 303-308.

[47] Chefd'hotel, C., Sebbane, A., (2007). "Random Walk and Front Propagation on Watershed Adjacency Graphs for Multilabel Image Segmentation," In: *Proc. of the Intern. Conference on Computer Vision.*
[48] Stawiaski, J., Decenciere, E., 2008. "Region merging via graph-cuts," *Image Analysis and Stereology*, 27(1), 39–45.
[49] Hanbury, A., 2008. "How do superpixels affect image segmentation?", *Proc.the 13th Iberoamerican Congress on Pattern Recognition (CIARP)*, Lecture Notes in Computer Science, 5197, 178–186.
[50] Zahn, C. T., "Graph-theoretic methods for detecting and describing gestalt clusters," IEEE Transactions on Computers, 20(1), 68-86 (1971).
[51] Horowitz, S.L., Pavlidis, T., "Picture segmentation by a tree traversal algorithm," Journal of the Association for Computing Machinery, 23(2), 368-388 (1976).
[52] Morris, O. J., de J. Lee, M., Constantinides, A. G., 1986. Graph theory for image analysis: an approach based on the shortest spanning tree, IEE Proc. vol. 133, No. 2, 146-152.
[53] Kruskal, J.B.J., "On the shortest spanning subtree of a graph and the traveling salesman problem," Proc. Am. Math. Soc., 7, 48-50 (1956).
[54] Montanvert, A., Meer, P., Rosenfeld, A., "Hierarchical image analysis using irregular tessellations," IEEE Trans. on Pattern Analysis and Machine Intelligence, 13(4), 307-316 (1991).
[55] Jolion, J.-M., Montanvert, A., "The adaptive pyramid, a framework for 2D image analysis," CVGIP: Image Understanding, 55(3), 339-348 (1992).
[56] Xu, Y., Uberbacher, E. C., "2D image segmentation using minimum spanning trees," Image and Vision Computing, 15, 47-57 (1997).
[57] Felzenszwalb, P.F., Huttenlocher, D.P., Image segmentation using local variation, Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition, 98-104 (1998).
[58] Haxhimusa, Y., Kropatsch, W. G., "Segmentation graph hierarchies," Lecture Notes in Computer Science, 3138, 343-351 (2004).
[59] Haxhimusa, Y., Ion, A., Kropatsch, W.G., "Comparing hierarchies of segmentations: Humans, normalized cut, and minimum spanning tree," In: Digital Imaging and Pattern Recognition, 95-103 (2006).
[60] Kropatsch, W.G., Haxhimusa, Y., Ion, A., "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 (2007).
[61] Borůvka, O. 1926. O jistém problému minimálnim. (about a certain minimal problem) Práce Moravské Přírodvédecké Společnosti v Brně (Acta Societ. Scienc. Natur. Moravicae), 3(3), 37-58 (In Czech.) (1926).
[62] Nešetřil, J., Milková, E., Nešetřilová, H., (2001). "Otakar Borůvka on minimum spanning tree problem. Translation of both the 1926 papers, comments, history," *Discrete Mathematics*, 233, 3-36.
[63] Prim, R.C., "Shortest connection networks and some generalizations," The Bell System Technical Journal, 36, 1389-1401 (1957).
[64] Canny, J., "A computational approach to edge detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6), 679-698, (1986).
[65] Shewchuk, J.R., "Triangle: engineering a 2D quality mesh generator and Delaunay triangulator," *Lecture Notes in Computer Science*, 1148, 203-222, (1996).
[66] Skurikhin, A.N., Volegov, P.L., "Object-oriented hierarchical image vectorization," In: Proceedings of the Int. Conference on Geographic Object Based Image Analysis (GEOBIA 2008) (2008).
[67] Skurikhin, A.N., "Proximity graphs based multi-scale image segmentation," Proceedings of the 4th Int. Symposium on Visual Computing, Lecture Notes in Computer Science, 5358, 298-307 (2008).
[68] Chica-Olmo, M., Abarca-Hernandez, F., "Variogram derived image texture for classifying remotely sensed images," In: *Remote Sensing Image Analysis: Including the Spatial Domain*, Springer, pp. 93-111, (2004).