

DISCRETE TOPOLOGY BASED HIERARCHICAL SEGMENTATION FOR EFFICIENT OBJECT-BASED IMAGE ANALYSIS: APPLICATION TO OBJECT DETECTION IN HIGH RESOLUTION SATELLITE IMAGES

Abdul Haleem Syed^{a,*}, Eli Saber^{a,b}, and David Messinger^a

^a Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, New York, USA 14623
– has1061@rit.edu, messinger@cis.rit.edu

^b Department of Electrical and Microelectronic Engineering, Rochester Institute of Technology, Rochester, New York,
USA 14623 – essee@rit.edu

Commission VI, WG VI/4

KEY WORDS: Multi-scale, Scale-space, segmentation, High-resolution, topological information, contextual information, region merging, SCRM, OBIA, GEOBIA, object detection

ABSTRACT:

With rapid developments in satellite and sensor technologies, there has been a dramatic increase in the availability of high resolution (HR) remotely sensed images. Hence, the ability to collect images remotely is expected to far exceed our capacity to analyse these images manually. Consequently, techniques that can handle large volumes of data are urgently needed. In many of today's multi-scale techniques the underlying representation of objects is still pixel-based, i.e. object entities are still described/accessed via pixel-based descriptors, thereby creating a bottleneck when processing large volumes of data. Also, these techniques do not yet leverage the topological and contextual information present in the image. We propose a framework for Discrete Topology based hierarchical segmentation, addressing both the algorithms and data structures that will be required. The framework consists of three components: 1) Conversion to dart-based representation, 2) Size-Constrained-Region Merging to generate multiple segmentations, and 3) Update of two sparse arrays SIGMA and LAMBDA which together encode the topology of each region in the hierarchy. The results of our representation are demonstrated both on a synthetic and a real high resolution images. Application of this representation to object-detection is also discussed.

1. INTRODUCTION

Over the past few years, there have been significant improvements in our ability to capture high-resolution satellite images. For instance, the recent WorldView-2 sensor can capture images at < 0.5 m resolution with a collection capacity of 300,000 sq mi/day. At this rate, this instrument alone can cover the entire USA in 12 days. Further, over this decade, it is projected that 288 earth observation satellites from 42 countries are to be launched (Euroconsult, 2012). Our ability to collect high-resolution data far exceeds our capacity to analyse them manually. Consequently, techniques for automated production of geospatial information and assisted image analysis are urgently needed.

To deal with segmentation of high-resolution remotely sensed images - one of the core tasks of image analysis - , a plethora of techniques have surfaced in the literature (Dey et al., 2010). The most sought after techniques are those that incorporate multi-resolution models through the use of appropriate scale space representations (Dey et al., 2010). This is because in complex high-resolution images, an object of interest to an analyst may reveal itself at any size/scale of observation; therefore analysis at multiple scales is necessary. Although these methods contribute significantly in dealing with complex high-resolution images (Baatz and Schape, 2000, Chen et al., 2009, and Syed et al., 2011), the underlying representation and information processing is still primarily pixel-based thereby creating a bottleneck when processing large volumes of data. Further, these methods do not fully leverage the topological information present in the images in order to improve detection results. The

ability to leverage contextual information requires examining a regions (potential object's) neighbourhood, and exploring the arrangement of adjacent regions (Syed et al., 2012). For instance, (Inglada and Michel, 2009) demonstrated successful use of spatial reasoning techniques to quantify topological relationships between image objects in an object detection algorithm using their multi-scale segmentation. However, their method of topological information extraction involved pixel-based processing which significantly limits real-time topological queries between any two regions. To mitigate the above mentioned problems, a discrete topology based framework for topological information extraction was proposed in (Syed et al., 2012). However, no framework was provided in reference to using this model for multi-level topological queries.

To overcome the above mentioned issues, we propose a framework for a discrete topology based multi-scale segmentation of high resolution satellite images. This proposed representation builds-on and improves our previous research on scale-space representation (Syed et al., 2011) and topological information extraction and encoding (Syed et al., 2012). Our goal is to provide an effective foundation/framework that will facilitate/assist analysts in tasks such as target detection/recognition, classification, change detection, and multi-sensor information fusion.

The remainder of this paper is organized as follows. Section 2, provides a description of the methods used in our framework. Section 3, presents the results of applying this framework to high resolution satellite images. Finally, Section 4 presents the conclusion.

2. FRAMEWORK FOR DISCRETE TOPOLOGY BASED HIERARCHICAL SEGMENTATION

The three main components of our framework are: 1) Conversion to dart-based representation, 2) Size-Constrained-Region Merging to generate multiple segmentations, and 3) Update of two sparse arrays SIGMA and LAMBDA which together encode the topology of each region in the multi-scale segmentation. Steps 2 and 3 take place in tandem as each level of the hierarchy is created. A block diagram of the proposed method is shown in Figure 1.

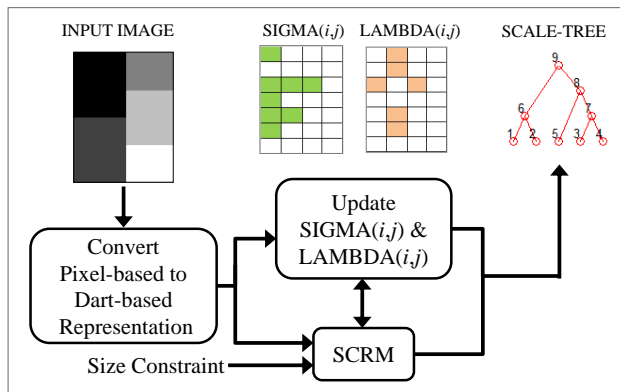


Figure 1: Block diagram of overall framework.

2.1 Conversion from Pixels to Darts

The first component of the process is designed to generate a dart-based representation from the pixel-based region map. This step reduces the dimensionality of the data significantly by representing regions by their *darts* (reduced boundary representation) thereby completely encoding the topology of the map. The resulting dart-based representation is a combinatorial map G such that $G = (D, \sigma)$ where D is the set of darts and σ is defined on D such that a cycle of the permutation σ denoted by σ^* encodes the nodes. The permutation σ for the map shown in Figure 2(b) is the first column of the array shown in Figure 6(a). For a more detailed explanation of how σ encodes the topology of the regions please see reference (Syed et al., 2012).

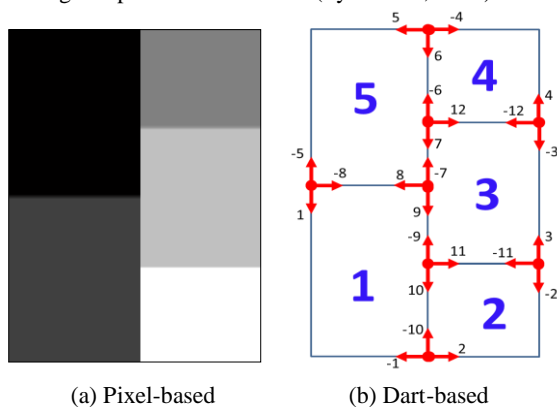


Figure 2: Conversion from pixel-based representation of regions to a dart-based representation.

The simple image in Figure 2(a) is of size 20x10, therefore uses 200 pixels to describe the five regions contained within. Compare this to dart-based representation, shown in Figure 2(b), which uses only 24 darts that not only describe the five regions, but also encode information for topological inferences.

2.2 Merging Mechanism to Generate a Hierarchy

The second component is the region merging mechanism to generate multiple segmentations. In this framework, the Size-Constrained Region Merging (SCRM) Algorithm (Castilla et al., 2008) is used to generate a hierarchy of segmentations by controlling the size of the objects that appear at any given level. We have adapted the SCRM to generate segmentations at multiple scales but the core idea of the algorithm remains the same. A summary of the algorithm is provided in Figure 3.

Algorithm 1: Size Constrained Region Merging

Input: • Size constraint for the given level max_sz
• list of regions at a given level $lbi[]$
Output: • list of regions at the next level $lbj[]$
• update of data structures

Find most similar neighbour

- 1 **for** each region in $lbi[]$
- 2 • find all its neighbours
- 3 • find its most similar neighbour $msn()$, based on spectral similarity

Iterative Region merging step

- 4 **for** each region in $lbi[]$
- 5 **if** $merge_condition$ based on size is *true* **then**
- 6 • merge region with its $msn()$
- 7 • Update new region properties and $SIGMA(i,j)$ and $LAMBDA(i,j)$
- 8 • Enforce merging constraints on neighbours of merged regions

Figure 3: Algorithm for SCRM

The overall algorithm is based on iterative merging of regions, uniformly across the image, until all regions below a given size (area) are eliminated. The $merge_condition$ in step 5 of the algorithm checks to ensure that the region being merged is smaller than the size constraint for that level and also ensures that this region and its most similar neighbour have not previously been merged during the same iteration. Enforcement of merging constraints in step 8 of the algorithm allows for controlled aggregation so that the resulting regions have the highest possible homogeneity given the size constraint. Which means that homogeneous regions are formed first, and then dissimilar gaps smaller than the size constraint are progressively incorporated into the former (Castilla et al., 2008). For more details on how the SCRM algorithm is used to generate a multi-scale segmentation, refer to (Syed et al., 2011).

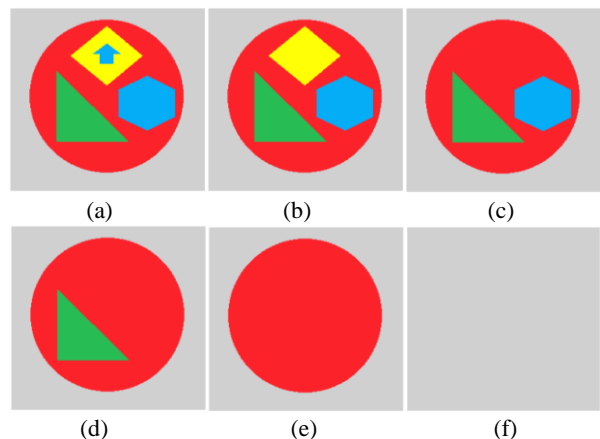


Figure 4: Application of the SCRM algorithm to a simple image shows how segmentations at multiples scales can be generated by altering the size constraint.

2.3 Data Structures & Procedures to Encode the Hierarchy

Finally, the third component is the procedure used to update the data structures that store the hierarchical scale-tree. As we deal with bigger and bigger remotely sensed images (area covered), the storage and efficiency becomes more important. Using a dart-based representation for each level, not only saves memory, but also enables efficient access and retrieval of the desired information from the hierarchy. The update of the data structures is performed simultaneously as the scale-tree is being built from one level to the next (Section 2.2).

The sparse arrays SIGMA and LAMBDA together encode the hierarchy of the scale-tree. They are similar in function to their one-dimensional counterparts described in (Syed et al., 2012). The columns of the SIGMA array are σ -permutations which capture the incidence relations between the darts, while the columns of the LAMBDA array capture the region labels associated with each dart. Every column of these structures represents a level in the hierarchy. Only those rows of the new column (level) which are different than rows of the previous column (level) are stored in the array making for an efficient storage of the hierarchy.

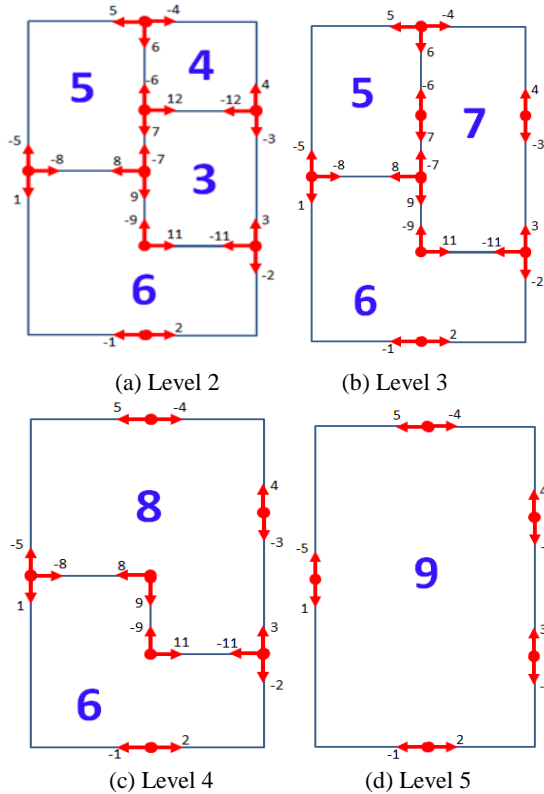


Figure 5: Multi-level segmentation of image from Figure 2(a). Note that progressively lesser number of darts will be needed to describe the regions as the levels go up.

For a given level, the permutation sigma ' σ ' and region labels ' λ ' can be extracted from the structures by using equations (1) and (2), where ' d ' represents the dart label and ' i ' represents the level in the hierarchy.

$$\sigma_i = \text{SIGMA}(d, i) \quad (1)$$

$$\lambda_i = \text{LAMBDA}(d, i) \quad (2)$$

For illustration, consider a hierarchy of regions generated for the image shown in Figure 2(a). The hierarchy contains five

levels with the first level shown in Figure 2(b) and the remaining four levels are shown in Figure 5. The corresponding SIGMA and LAMBDA arrays which encode these five levels are shown in Figure 6. Note that as the number of levels increase, the columns of the arrays become sparser. Removal of a region at a given level involves removing the darts associated with that region and updating the data structures SIGMA and LAMBDA. The steps used to update each of these structures will be briefly described in the proceeding sections.

SIGMA($dart, level$)						LAMBDA($dart, level$)					
Darts	Level 1	Level 2	Level 3	Level 4	Level 5	Darts	Level 1	Level 2	Level 3	Level 4	Level 5
1	-8	0	0	0	-5	1	0	0	0	0	0
-1	2	0	0	0	0	-1	1	6	6	6	9
2	-10	-1	0	0	0	2	0	0	0	0	0
-2	3	0	0	0	0	-2	2	6	6	6	9
3	-11	0	0	0	-2	3	0	0	0	0	0
-3	4	0	0	0	0	-3	3	3	7	8	9
4	-12	0	-3	0	0	4	0	0	0	0	0
-4	5	0	0	0	0	-4	4	4	7	8	9
5	6	0	0	-4	0	5	0	0	0	0	0
-5	1	0	0	0	0	-5	5	5	5	8	9
6	-4	0	0	0	0	6	5	5	5	0	0
-6	7	0	0	0	0	-6	4	4	7	0	0
7	12	0	-6	0	0	7	5	5	5	0	0
-7	8	0	0	0	0	-7	3	3	7	0	0
8	9	0	0	0	0	8	5	5	5	5	0
-8	-5	0	0	0	0	-8	1	6	6	6	0
9	-7	0	0	8	0	9	1	6	6	6	0
-9	10	11	0	0	0	-9	3	3	7	7	0
10	11	0	0	0	0	10	1	0	0	0	0
-10	-1	0	0	0	0	-10	2	0	0	0	0
11	-9	0	0	0	0	11	2	0	0	0	0
-11	-2	0	0	0	0	-11	3	3	7	8	0
12	-6	0	0	0	0	12	3	3	0	0	0
-12	-3	0	0	0	0	-12	4	4	0	0	0

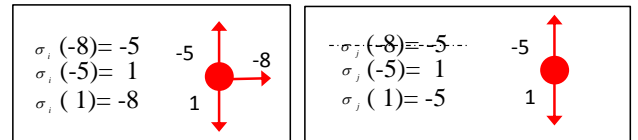
(a) SIGMA (Σ)

(b) LAMBDA (Λ)

Figure 6: Recall that the arrays sigma ' σ ' and lambda ' λ ' encode the topology of a single segmentation (Syed et al., 2012). SIGMA and LAMBDA are their multi-dimensional equivalents

2.3.1 Update of SIGMA

Recall that the permutation σ is stored as an integer array which encodes the dart incidence relationships in the counter-clockwise direction (Syed et al., 2012), as shown in Figure 7(a).



(a) σ_i for a node at level i

(b) σ_j for a node at the next level j

Figure 7: Encoding and updating of σ from one level to next.

When using the dart-based representation of regions, the region merging operation consists of dart removal. Figure 7 illustrates the removal of dart -8 and the updated σ relationships. A detailed description of removal operation (Brun et al., 2003) is beyond the scope of this paper, but the elementary algorithm is shown in Figure 8.

Algorithm 2: Updating the permutation σ

Input: • Array σ_i
• darts to be removed $K = \{d_k\}$
Output: • Array σ_j
Update the incidence relationships in σ_j

- 1 **for** each dart $d_k \in K$
- 2 • set $\sigma_j(d_k) = \text{null}$
- 3 • set $\sigma_j(\sigma_i^{-1}(d_k)) = \sigma_i(d_k)$
- 4 • set $\sigma_j(\sigma_i(d_k)) = \sigma_i^{-1}(d_k)$

Figure 8: Basic algorithm to update permutation σ

Consider the merging of regions R6 and R8 from Figure 5(c) to get region R9 in Figure 5(d). The update procedure involves removing the dart set $K = \{-8, 8, 9, -9, 11, -11\}$. By comparing the columns Level4 and Level5 of Figure 6(a), notice the updates that were performed using Algorithm 2 i.e. $\sigma_4(1) = -8$ (note that $\sigma_4(1) = 0$ in SIGMA(1,4) but its value can be inferred from its previous non-zero value in the table) has been updated to $\sigma_5(1) = -5$ and $\sigma_4(3) = -11$ has been updated to $\sigma_5(3) = -2$.

2.3.2 Update of LAMBDA

As a new region is created at a level, its associated darts are labelled appropriately and stored in the array LAMBDA. Let R_l be the result of merging regions R_i and R_j , then the algorithm to update the dart associated region labels is shown below in Figure 9.

Algorithm 3: Updating the λ array for each level

Input: • darts $d_i \in R_i$ and $d_j \in R_j$

• removal dart set $d_k \in K$

Output: • darts set defining $R_l = \{d_l\}$

- 1 **for** each dart $d_k \in K$
- 2 • set $\lambda(d_k) = \text{null}$
- 3 **for** each dart $d_s \in \{(d_i \cup d_j) - d_k\}$
- 4 • set $\lambda(d_s) = l$
i.e. new label assigned to surviving darts

Figure 9: Basic algorithm to update λ

Once again, consider the merging of regions R6 and R8 in Figure 5(c) to get region R9 in Figure 5(d). The update procedure involves removing the dart set $\{-8, 8, 9, -9, 11, -11\}$. Note that the LAMBDA entries in the last column of Figure 6(b), which correspond to the removal darts, have been set to zero. The surviving darts $\{-1, -5, -4, -3, -2\}$ that define region R9 have been given their proper label.

3. RESULTS

Using the framework described in Section 2, a hierarchical scale-space representation of the images is generated. The encoding of this hierarchy using LAMBDA and SIGMA allows us to efficiently reconstruct the multiple segmentations for further processing that may be required. A visualization of the results of this representation is shown for both synthesized and real images. Note, that for this visualization, each region is represented by a node in the tree. The position of the node on the x-y plane is the centroid of the region. The position of the node on the vertical axis is a function of its position in the scale-space.

3.1 Visualization of the Results

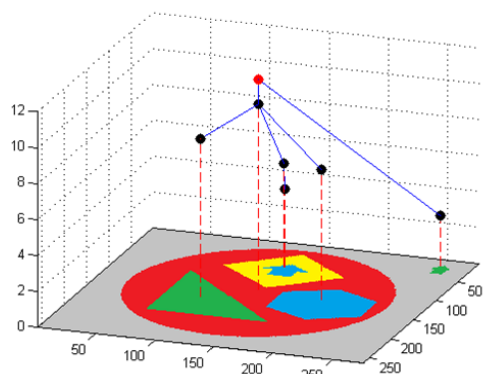


Figure 10: Scale-Tree of image shown in Figure 4.

3.2 Results of Using Hierarchical Representation on High Resolution Satellite Images

This image is a small section from Le Faux, France image, courtesy Digital Globe©. Figure 11(a) through (f) shows a few levels of the multi-scale segmentations that are stored within the hierarchy. In Figure 11(f), note how the man-made objects (houses and roads) have been separated from the natural parts of the image scene (trees and grass), such a delineation is not readily apparent in Figure 11(a). As a result, this hierarchical representation allows an image analyst to observe and exploit information in the image that occurs at multiple scales. Figure 11(g) is a scale-tree visualization of the image content where each region is represented as node in the tree, positioned directly above the centroid of the region.

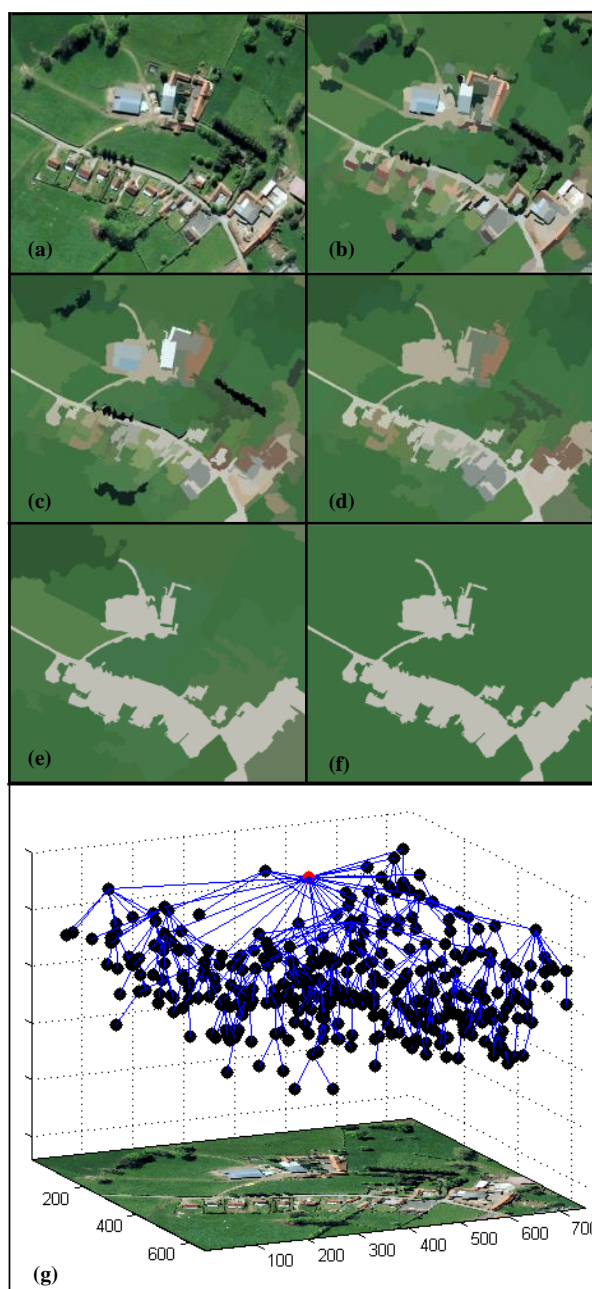


Figure 11: (a)-(f) A few levels contained in the multi-scale hierarchy of a high resolution image (g) The scale-tree of the image. Scene WorldView-2 sensor courtesy DigitalGlobe©

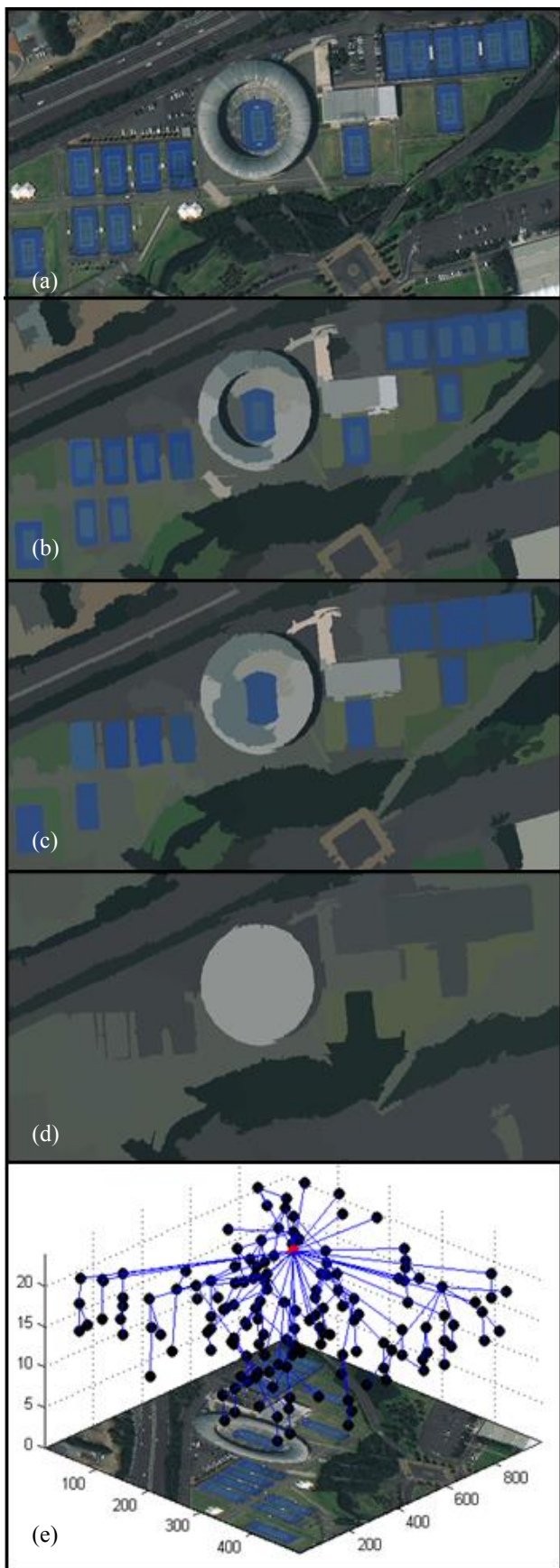


Figure 12: (a)-(f) A few levels contained in the multi-scale hierarchy of a high resolution image (g) The scale-tree of the image. Scene WorldView-2 sensor courtesy DigitalGlobe© (Image : Sydney Olympic Complex, Australia)

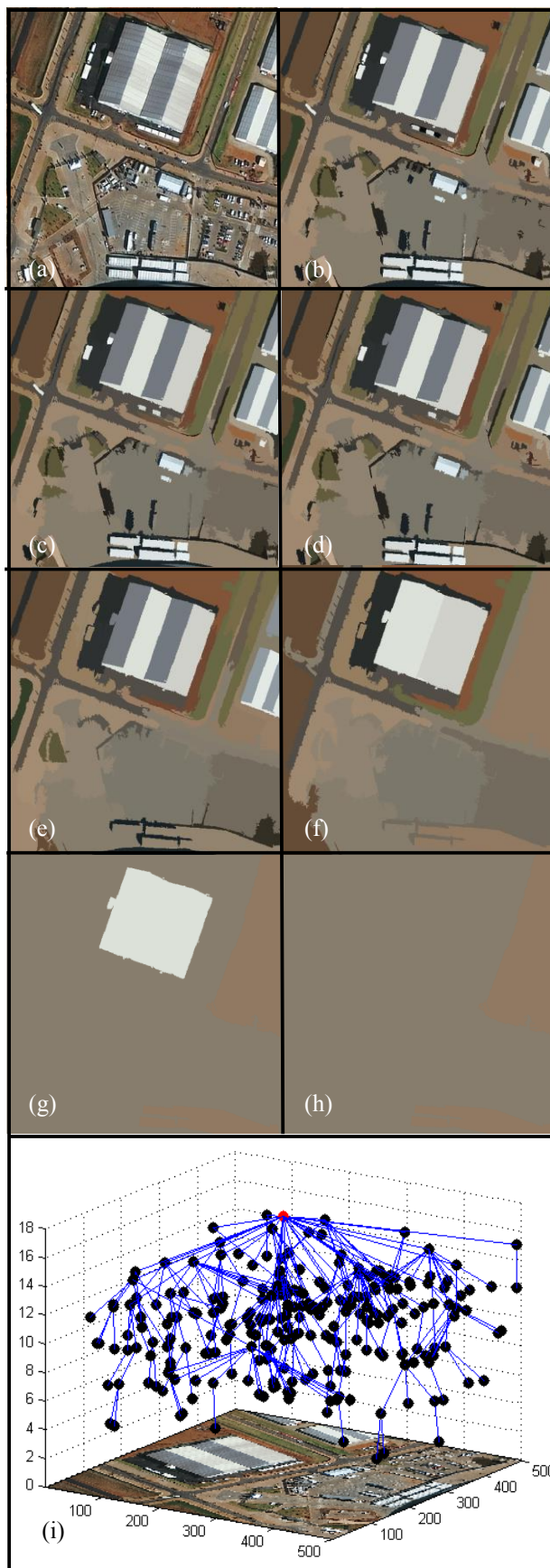


Figure 13: (a)-(h) A few levels contained in the multi-scale hierarchy of a high resolution image (i) The scale-tree of the image. Scene WorldView-2 sensor courtesy DigitalGlobe© (Image : Soccercity, Johannesburg, South Africa)

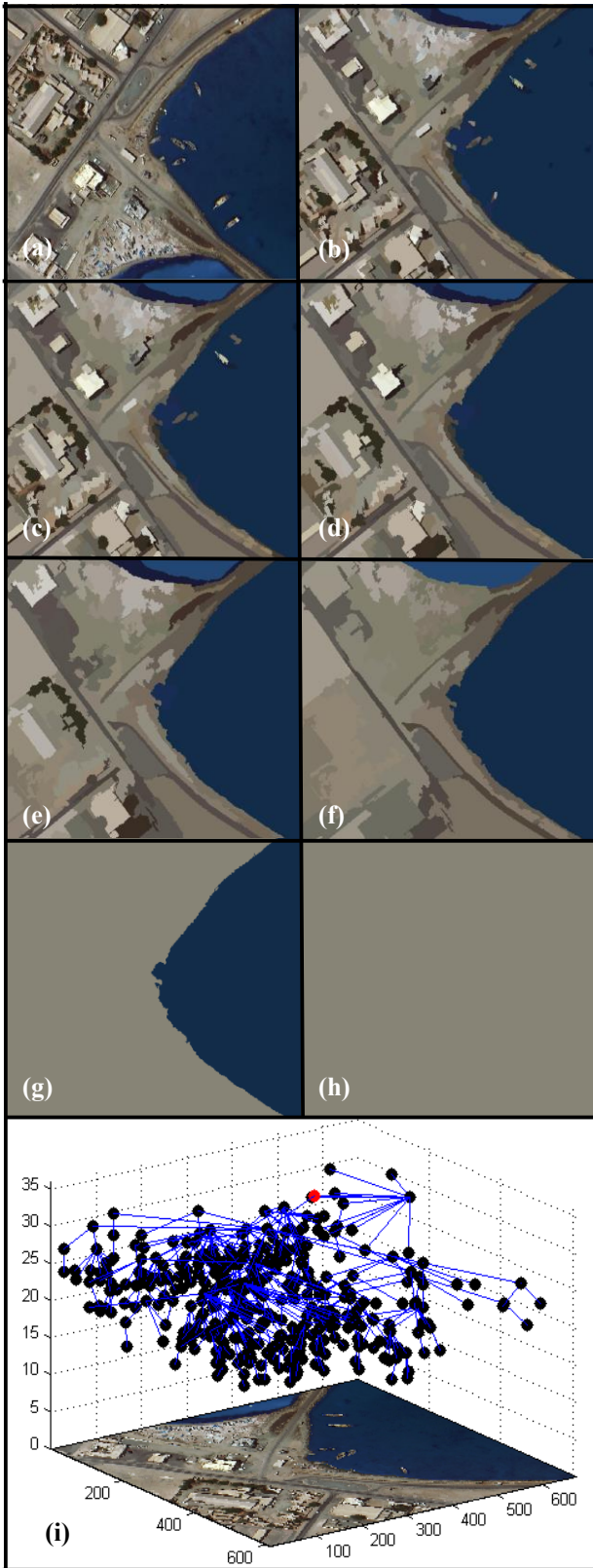


Figure 14: (a)-(h) A few levels contained in the multi-scale hierarchy of a high resolution image (i) The scale-tree of the image. Scene WorldView-2 sensor courtesy DigitalGlobe© (Image : Al-Masirah Island Oman)

4. CONCLUSIONS

A framework for hierarchical segmentation using discrete topology was presented. First, the pixel-based multi-spectral

image was converted to a dart-based representation. Then, the data structures and algorithms required to create and store the hierarchical representation were shown. SCRM algorithm was used to generate multiple segmentations of the hierarchy by controlling the size constraint parameter.

Application of this representation for four high resolution images from the WorldView-2 Sensor, was illustrated in the results section. Note that that this representation was generated completely in an unsupervised fashion, and yet provides the image analyst a useful tool to sort and navigate the multiple segmentations of a scene. Since the image has been converted into a tree, the next stage of this research involves automatic detection of objects of interest, through attributed tree matching techniques (Kriege et al., 2012).

5. ACKNOWLEDGEMENTS

The authors would like to thank DigitalGlobe© for providing us with the high resolution images used in this paper, which were captured by the WorldView-2 Sensor.

6. REFERENCES

- Baatz, M., and Schape, A., 2000, "Multiresolution Segmentation: An Optimization Approach for High Quality Multi-scale Segmentation", *Angewandte Geographische Informationsverarbeitung*.
- Brun, L., Domenger, J.P., and Mokhtari, M., 2003, "Incremental Modifications of Segmented Image by Discrete Maps", *JVCIR* 14(2003) 251-290
- Castilla, G., Hay, G., and Ruiz-Gallardo, J., 2008, "Size Constrained Region Merging: An Automated Delineation Tool for Assisted Photointerpretation", *ASPR* 74 (4).
- Chen, J., Pan, D., and Maz, Z., 2009, "Image-object detectable in multiscale analysis on high-resolution remotely sensed imagery", *IJRS*.
- Dey, V., Zhang, Y., and Zhong, M., 2010, "A review on image segmentation techniques with remote sensing perspective," *ISPRS TC VII Symposium XXXVIII*.
- Euroconsult Report 2012, "Satellite-Based Earth Observation: Market Prospects to 2021." <http://www.euroconsult-ec.com/report/download?rep=contents&name=satellite-based-earth-observation-2012-toc.42.pdf> (1st Feb, 2013)
- Inglada, J., and Michel, J., 2009 "Qualitative spatial reasoning for high-resolution remote sensing image analysis," *TGRS*.
- Kriege, N., and Mutzel, P., 2012, "Subgraph Matching Kernels for Attributed Graphs", *ICML*
- Syed, A., Saber, E. and Messinger, D., 2011, "Scale-space Representation of Remote Sensing Images using an Object-oriented Approach", *Vol. 8053 (1), SPIE*.
- Syed, A., Saber, E. and Messinger, D., 2012, "Encoding of Topological Information in Multi-Scale Remotely Sensed Data: Applications to Segmentation and Object-Based Image Analysis", *GEOBIA*.