South African Journal of Geomatics, Vol. 2, No. 4, August 2013

A Technique for Optimal Selection of Segmentation Scale Parameters for Object-oriented Classification of Urban Scenes

Guy Blanchard Ikokou, Julian Smit

Geomatics Division, University of Cape Town, Rondebosch, South Africa, ikokou@yahoo.fr

Abstract

Multi-scale image segmentation produces high level object features at more than one level, compared to single scale segmentation. Objects generated from this type of segmentation hold additional attributes such as mean values per spectral band, distances to neighbouring objects, size, and texture, as well as shape characteristics. However, the accuracy of these high level features depends on the choice of segmentation scale parameters. Several studies have investigated techniques for scale parameter selection. These proposed approaches do not consider the different objects' size variability found in complex scenes such as urban scene as they rely upon arbitrary object size measures, introducing instability errors when computing image variances. A technique to select optimal segmentation scale parameters based on image variance and spatial autocorrelation is presented in this paper. Optimal scales satisfy simultaneously the conditions of low object internal variance and high inter-segments spatial autocorrelation. Applied on three Cape Town urban scenes, the technique produced visually promising results that would improve object extraction over urban areas.

Key words: segmentation, object oriented classification, object's variance, spatial autocorrelation, objective function, Moran's index.

1. Introduction

Multi-scale image segmentation is the partition of an image into spatially continuous, mutually disconnected and homogeneous regions at various segmentation levels (Pekkarinen, 2002). The technique was initially investigated by Woodcock and Strahler (1987). In the context of image analysis, scale is defined as the level of aggregation and abstraction at which an object can be clearly described (Benz *et al*., 2004). Multi-scale segmentation starts considering each pixel as an object and merges them to create larger objects based on homogeneity thresholds defined by the analyst (Blaschke, 2010). Multi-scale segmentation provides high level object features, compared to single scale segmentation. In fact, objects generated from this type of segmentation hold additional attributes such as mean value per spectral channel, distances to neighbouring objects, size, as well as shape characteristics. Multi-scale segmentation has the advantage of considering homogeneity criteria such as colour, shape compactness and smoothness, during the creation of image objects. The technique offers the possibility of varying the size of output segments and creates object hierarchy levels that facilitate their accurate extraction. In the absence of objects hierarchy levels, each object will be created from scratch and no topology relationships will be built between objects

produced at finer scales and those generated at coarser resolutions. Creating image hierarchy allows each segmentation level, except the first one, to be derived from previous levels. However, the quality of objects created from segmentations, relies on the choice of scale parameter, and this choice mostly relies on subjective series of trial and error (Hay *et al*., 2003; Meinel and Neubert, 2004; Kim *et al*., 2008). This paper proposes a technique that improved the selection of optimal segmentation scales by revealing thresholds that produce segments with low internal and high inter-segment variability for urban scenes.

2. Materials and Methods

2.1 Material

Three different scenes of Cape Town urban area, in South Africa, were considered to test the technique (Figure 1). The areas were extracted from a 0.5m spatial resolution ortho-rectified RGB digital aerial photograph. The choice of these areas was motivated by the fact that each scene was characterized by a different type of urban land cover, ranging from small sized buildings to vegetation extents.

Figure 1A: Original Cape Town urban scene. B: Residential area of Kensington composed of individual housing units of small and medium size, a road network, trees and recreation areas. C and D: Residential areas of Vredehoek and Oranjezicht.

The merit of segmentation results is influenced by the type of land cover and the number of spectral signature considered. This investigation used 1464 segmented objects which provided 1464 distinct spectral signatures samples and different segment size measures to calculate image internal and inter-segment variances. The different urban scenes were partitioned beforehand into several segmentation levels using the multi-resolution algorithm in the eCognition software. Once

segmented, the different objects' spectral and size statistics were used to determine segments variances and spatial autocorrelations.

2.2 Image segmentation

The different segmentations were done using 12 distinct scale parameters in the eCognition software, using the multi-resolution algorithm. Multi-resolution algorithm starts with one seed pixel object and merges each pair wisely with neighbouring objects to form larger objects until the homogeneity thresholds defined are reached. These homogeneity thresholds are controlled by the scale parameter which influences the size of output segments. The different image bands which are very relevant in terms of spectral information were equally weighted at a factor of 1 in order to consider all the spectral information made available by each band. The remaining parameters including shape, colour, compactness, smoothness were uniformly weighted at a factor of 0.5 for equal influence on the results.

2.3 Identifying optimal segmentation scales

The technique investigated in this research is comparable to those of Espindola *et al.*, (2006), Kim *et al.*, (2008), Johnson and Xie, (2011). The approach used internal and inter-segment heterogeneity to evaluate local and global segmentation quality. Two properties characterizing good segmentation were considered: (1) resulting segments must internally be homogeneous and (2) each segment should be separable from its neighbours. As a consequence, optimal scales should satisfy low internal and high intra-segment variances (each region must be homogeneous and adjacent regions should be dissimilar). The internal segment variance expressing the overall homogeneity of image objects was calculated using equation [1]:

$$\text{variance} = \frac{\sum_{i=1}^{n} a_i v_i}{\sum_{i=1}^{n} a_i} \qquad [1]$$

In equation [1], v_i is the variance and a_i is the area of the segment *i.* The internal object variance was weighted with the area size in order to put more weight on larger objects and limit possible instabilities caused by smaller ones. The inter-object variance expressing the spatial autocorrelation between neighbouring image objects was calculated using Moran's Index (MI) in equation [2]. This index was chosen because it is a reliable indicator of statistical separation between spatial objects (Johnson and Xie, 2011):

$$MI = \frac{n * \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} * (y_i - y)(y_j - y)}{(\sum_{i=1}^{n} (y_i - y)^2 (\sum_{i \neq j} \sum w_{ij}))} \qquad [2]$$

where (y_i -y) and (y_j -y) are the deviations from the means. W_{ij} is a measure of the spatial proximity of adjacent image objects. In this case a value of 1 was attributed to W_{ij} because image objects produced by the segmentation are adjacent objects (Espindola *et al.*, 2006).

The Moran's Index captures the difference between mean values of each image object and their neighbours. Low values of the index indicate high inter-object heterogeneity, which is an advantage for image segmentation. In order to equally consider the internal objects' variances and the Moran's measurements, the indices were normalised using equation [3], proposed in Espindola *et al.*, (2006):

$$X\,norm = \frac{X - X\min}{X\max - X\min} \quad [3]$$

with X taking values of image variance or the Moran's Index to be normalised.

Once the two statistical indicators have been normalised, they were combined to calculate the Objective Function given by equation [4]:

$$F(v, MI) = F(v) + F(MI) \quad [4]$$

where v is the image variance and MI is the Moran's Index. In order to determine the optimal scales, the different objective function values were plotted to produce a curve. The lowest values of the curve indicate potential optimal segmentation scales which produce segments with high spatial auto-correlation and low internal variance measures.

2.4 Refining the segmentation scales

After identifying the ideal scales, a refining function is applied to each value in order to measure the degree of internal homogeneity associated to each object. This function is defined by the equation [5] proposed in Johnson and Xie (2011):

$$Heterogeneity = \frac{normalisedVar - normalisedMI}{normalised\ \mathrm{var} + normalisedMI} \quad [5]$$

Low values of the index reveal low degree of objects' homogeneity and high values indicate more homogeneous image objects.

2.5. Accuracy assessment

The accuracy assessment of our technique was based on several reference objects (manually digitized in ArcGIS from the available aerial photography) and image objects segmented using the optimal scales identified in this research. In addition to visual analysis, the following three criteria were considered as indicators of the accuracy of the segmentation: (1) the percentage of the largest segmented object after excluding the 'extra' pixels which fall out of the reference object boundary, (2) the percentage of area of the 'lost' pixels and (3) the percentage of the area of the 'gained' or 'extra' pixels. If the area of the lost or extra pixels is more than 25% of the reference object, then there is a chance that the actual shape of the object is distorted by the segmentation process. In

addition, objects found with an area percentage more than 55% compared to the reference object are considered as optimally segmented (Marpu *et al.*, 2010).The value of 55% was chosen because the spectral value of the image can still be restored using the median value of the object reflectance to improve the results. Figure 2 illustrates the effects of gained and lost pixels.

Figure 2: An illustration of a manually digitised reference object overlaid with the segmented object. Gained or extra pixels and lost pixels mostly originate from spectral similarities with neighbouring objects. The areas of lost or gained pixels are characteristics of under segmentation and the percentage area of the biggest segmented object defines over segmentation.

2.6. Comparison with other segmentation techniques

The quality of segmentation results can be tested by comparing segment objects with reference objects using formal properties such as object difference. This object difference can be analysed as a geometric relationship between segments and its corresponding reference. In the case of a good segmentation, the similarity between the image object and its reference should be maximized. Four analyses were done to compare the segmentation results from our technique to other approaches: (1) the percentage of overlapping area between image objects and respective references, (2) the average difference of perimeter, (3) the Area-Fit-Index and the shape index and (4) the Shape Index.

The Area-Fit-Index proposed in Lucieer (2004) is defined in equation [6]:

$$AFI = \frac{Area\ \ reference - Area\ \ largest\ \ segment}{Area\ \ reference\ \ object} \qquad [6]$$

The shape index was calculated using the equation provided in Neubert and Herold (2008) as follows:

$$Shape\ Index = \frac{Perimeter}{4\sqrt{Area}} \quad [7]$$

where Perimeter and Area represent the image object's perimeter and area measures.

3. Experimental Results

3.1 Identifying suitable scales

In addition to visual interpretation, optimal segmentation scales can be evaluated using internal variance and spatial autocorrelation measures. Kim *et al.*, (2008) described optimal segmentation scales as being associated with negative Moran's Index values and the spatial statistics associated with the optimal scales found in this study were located below zero as shown in table 1.

Table1: Weighted and normalised variances and Moran's Indices associated with each scale parameter.

Scale parameters	Weighted variance	Moran's Index	Normalised variance	Normalised Moran's Index	Objective function
15	1.7934	-0.0786	0.7761	0.3491	1.1252
25	1.6487	1.1792	0.7309	0.9640	1.6949
35	1.5935	-0.3374	0.7535	0.2628	1.0163
55	1.6745	0.0262	0.6787	0.9421	1.6208
65	1.7018	1.0425	0.7450	0.9437	1.6887
85	1.6722	-0.3599	0.5446	0.8674	1.4120
100	1.7953	0.1224	0.8577	0.8406	1.6963
120	1.7922	-1.1504	0.5945	0.0248	0.6193
150	1.6733	1.1760	0.6397	0.9927	1.6324
180	1.6545	0.0345	1.1263	0.5533	1.6796
200	21.3007	-0.0062	0.5444	0.3439	0.8883
260	1410.9208	18.3254	0.8318	0.5479	1.3797

The scale parameter of 15 was chosen as our starting scale and as a consequence it was not considered as a low point of the objective function curve despite its negative Moran's Index value as revealed in table 1.

The objective function values calculated for the three study sites revealed in figure3 that scales of 35, 120, 120 and 200 are optimal for segmentation. The curve presented a localised normal distribution with succession of high and low points.

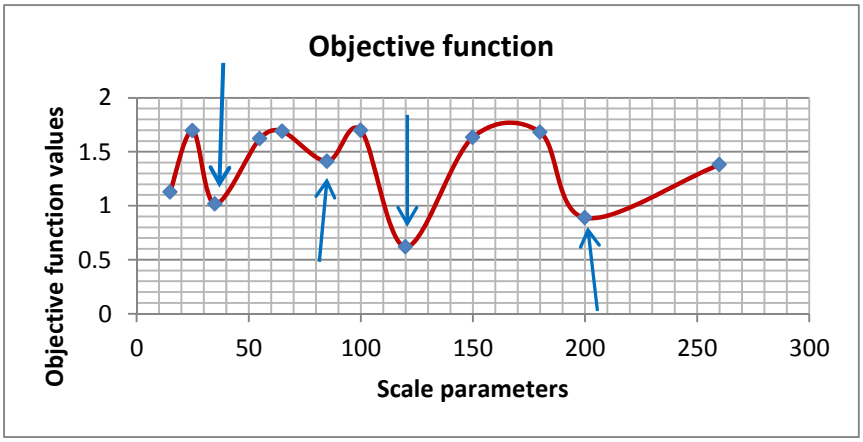

Figure3. Objective function showing six low values describing optimal segmentation scales. Scales of 35, 85,120 and 200 were identified.

3.2 Refining under and over segmentation

In order to refine the different optimal scales, a heterogeneity function was used. The curve of this index describes more homogeneous objects at the peak points. In contrast, lower points of the curve characterize objects with high internal heterogeneity. Figure 4 presents the different heterogeneity scores associated with each scale.

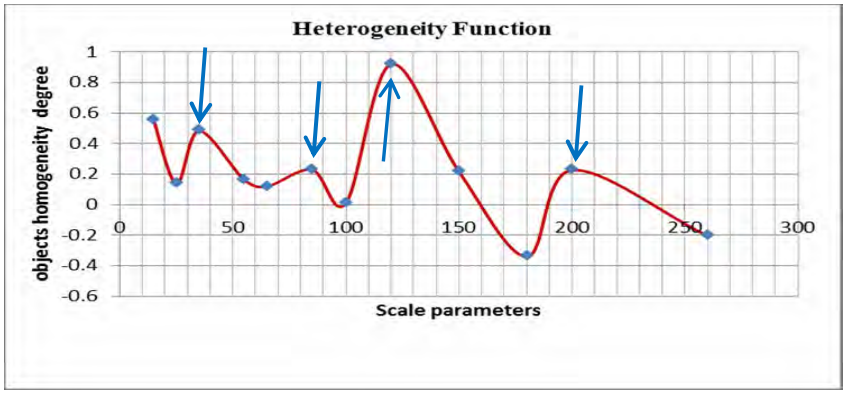

Figure 4. Heterogeneity function showing segments' internal homogeneity scores at various scales. The higher the value the more homogeneous is the segment. Scales of 120 and 200 for instance produced better segments homogeneity than 150 and 250.

Figure 5 illustrates visual segmentation results of two sport fields in the Kensington scene. As can be observed, more than 70% of the reference areas overlap with more than 85% of segments' areas. A good delineation of both sport fields was achieved at scale of 85. However, the effect of over segmentation was prominent at scales of 25, 55 and 65. At these scales segments' boundaries have not reached the reference objects' sizes.

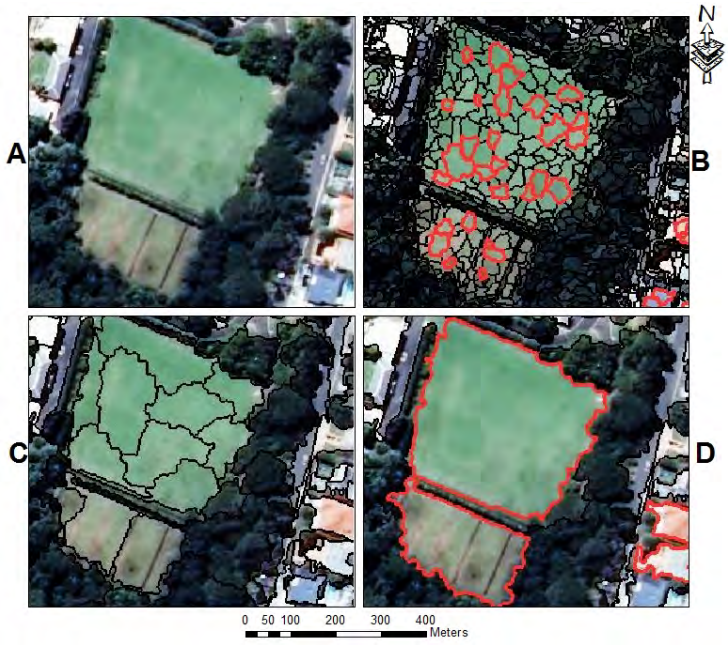

Figure 5: A: Original Kensington scene, 5B: segmentation at scale of 25 showing sport fields partitioned into smaller segments, 5C: over-segmentation at scale of 55 and 5D: good segmentation results at scale of 85.

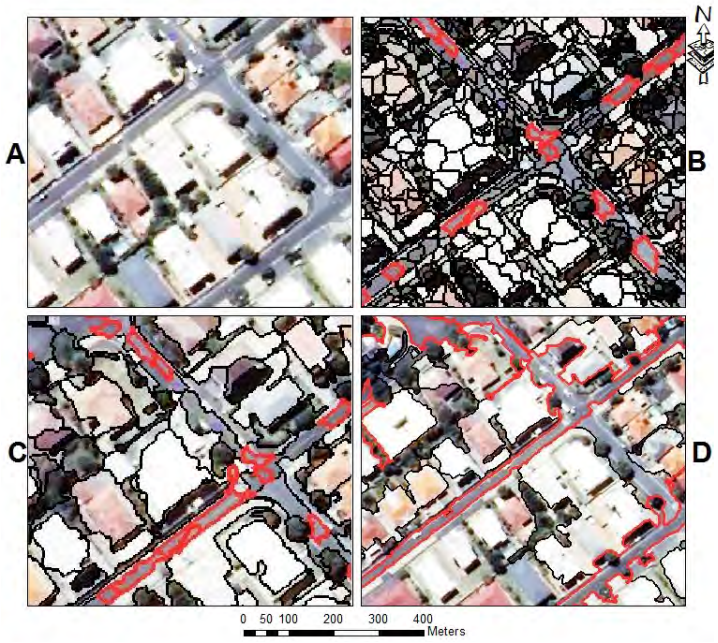

Figure: 6A: Original Oranjezicht scene, 6B and 6C: segmentations at respective scales of 25 and 55 showing small road segments and 6D: good segmentation of the same area at scale of 120 showing segments approximating their real world widths and lengths.

The segmentation performed on the Oranjezicht scene (see figure 6) produced a satisfactory delineation of the road network, when done at scale of 120. At lower scales, the process produced an over-segmentation with small individual road segments difficult to classify. The proposed scale

exhibited high discriminative power against road intra-class variations in an image containing trees and pavement areas. Moreover, more than 90% of road segments areas overlapped with the reference objects' boundaries at scale of 120 and the amount of lost pixels were very low in comparison to scales of 25 and 55. However, some gained pixels areas were observed mostly at road junctions due to the presence of large pavements. The scale parameters 25 and 55 described high objects' internal variances, explaining the partitioning of roads into smaller objects. Increasing the scale threshold up to 120 reduced this high internal variance, resulting in a better delineation of the road network as illustrated in figure 6D.

In figure 7, the scale parameter of 25 produced individual trees with size approximating some small buildings, limiting their separation from the built up structures. Forest stands, easy to extract based on size and shape index, were obtained at scale of 200 over the area of Vredehoek.

Figure: 7A: Original Vredehoek scene, 7B and 7C: segmentation at scales of 15 and 25 showing individual trees and 7D: segmentation showing group of trees easy to classify at scale of 200.

The visual analysis of segmentations done at scales of 120 and 150 revealed that the scale of 150 produced objects containing more than one class as illustrated in figure 8. In contrast, the scale of 120 minimized the effect of mixed pixels.

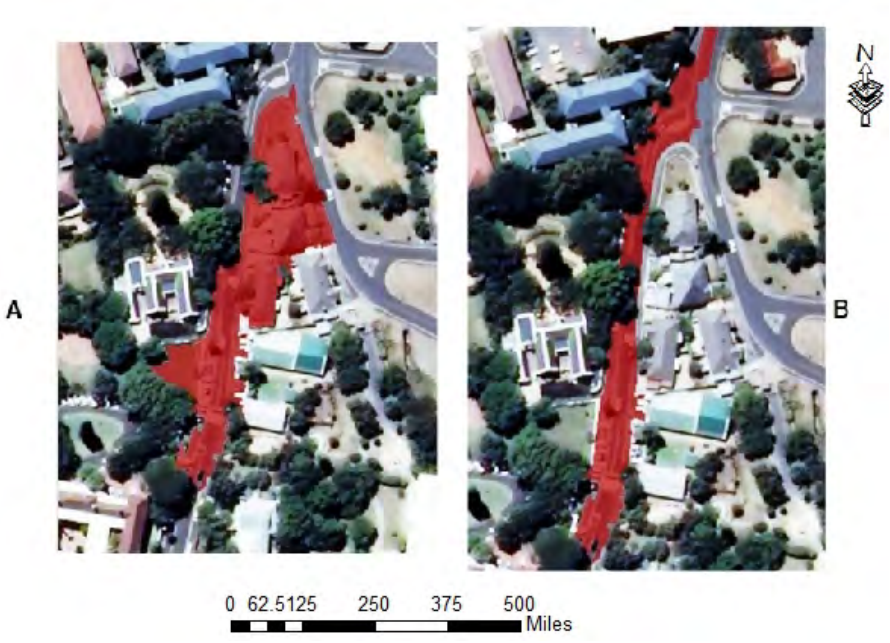

Figure 8A: segmentation at scale of 150 producing heterogeneous objects. 8B: segmentation done at scale of 120 optimally captured a portion of road in comparison to A.

3.3 Accuracy Assessment

Table 2 shows an extract of the accuracy assessment of 30 randomly selected segmented reference objects. The overall results show that the area overlap between objects and their references was above 80% and the different percentages of lost pixels were below 20%. In the same way, the percentages of gained pixels were found between 1 and 17 %.

Table 2: Area proportions of the largest segmented objects as well as the percentage of 'gained' (extra) and 'lost' pixels.

Object sample	Area reference	Area object	Percentage Segmented object	Percentage extra pixel	Percentage lost pixel
1	402.36	361.01	89.72%	0	10.28 %
2	293.17	303.15	126%	3.40%	0 %
3	467.11	472.17	101.08%	1%	0 %
4	483.56	396.06	81.91	0%	18.09 %

4. Discussion and Conclusion

To date the selection of segmentation parameters has mostly been performed intuitively by 'trial and error'. This paper presented an improved concept for the selection of optimal segmentation scales based on image variance and spatial auto-correlation. It extends the methods described in Espindola *et al.*, (2006), Kim *et al.*, (2008) and Johnson and Xie (2011) by considering real world objects' size measures for minimizing instabilities caused by arbitrary size thresholds. This research provides a comprehensive approach to understanding the segmentation results and also provides support for finding optimal segmentation scales. The segmentation evaluation was based on 30 manually digitized reference objects per image. The results of the evaluation of our technique

compared with other approaches presented in Neubert and Herold (2008) are shown in table 3. Based on the mean area difference, the technique exposed in this research demonstrated a better performance in balancing under and over-segmentation. This is illustrated by the smallest mean area difference value found in comparison to the other approaches. The ENVI 4.4 method also had a good performance in limiting under and over-segmentation followed by the EDISON segmentation approach. Comparing our segmentation with the EWS 1.0 segmentation, the latter resulted in high amount of lost and extra pixels illustrated by the high average area of 24.6. The ENVI 4.4 segmentation performed better in terms of object perimeter followed by our segmentation technique. The Berkley Imgseg 0.54 and InfoPACK 2.0 methods produced the worst results when it comes to object delineation.

Table 3: Comparison of various segmentation techniques based on 30 random reference objects.

Segmentation techniques	**ENVI 4.4**	**Berkley Imgseg 0.54**	**Definiens Developer 7.0**	**EDISON**	**EWS 1.0**	**InfoPACK 2.0**	**Our Technique**
Number of reference areas	20	20	20	20	20	20	**30**
Average Difference of area (m^2)	6.9	12.3	15.9	11.5	24.6	17.0	**5.4**
Average difference of perimeter (m)	12.3	22.2	17.2	13.8	18.1	29.6	**11.44**
Average difference of shape index	10.8	21.1	16.2	12.4	15.4	46.0	**7.29**
Area-Fit-Index	-0.04	-0.14	0.08	-0.18	-0.12	-0.04	**-0.19**

The differences in object shape between the references and segmented objects were minimized using our technique, followed by ENVI 4.4. The InfoPACK 2.0 segmentation produced highly distorted objects' shapes illustrated by the average shape index value of 46.0. All the techniques performed well regarding the Area-Fit-Index. The smaller the Area-Fit-Index value, the better is the segmentation.

This analysis confirms and supports our previous visual observation results. Despite some losses and gains of pixels that affected mostly roads and large buildings the proposed technique improved the segmentation quality by minimizing over and under-segmentations. This makes it a valuable source for selecting the best-fitting multi-level segmentation scales.

5. References.

Benz, UC Hofman, P Willhauck, G Lingenfelder, I and Heynen, M 2004, "Multi-resolution object-oriented fuzzy analysis of remote sensing data for GIS ready information", *Photogrammetry and Remote Sensing*, Vol.58, No.3/4, pp. 239-258.

Blaschke, T 2010, "Object-based image analysis for remote sensing" *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(2010) 216.

Dragut, L Tiede, D and Levisick, R 2010, "ESP: a tool to estimate scale parameter for multi resolution image segmentation of remotely sensed data", *International Journal of Geographical Information Science*, Vol. 24, No. 6, June 2010, 859-871.

Espindola, GM Camara, G Reis, I A Bins, LS and Montero, AM 2006, " Parameter Selection for Region-Growing Image Segmentation Algorithms using Spatial Autocorrelation" *International Journal of Remote Sensing*, in press, 2006.

Hay, G Blaschke,T Marceau and Bouchard, A 2003, "A comparison of three image object Methods for the Multi scale Analysis of Landscape Structure" *ISPRS Journal of Photogrammetry and Remote sensing*, Vol. 57, No.5-6, pp. 327-345.

Johnson, B and Xie, Z 2011, "Unsupervised image segmentation evaluation and refinement using multi-scale approach" *ISPR Journal of Photogrammetry and Remote Sensing* 66(2011) 473-483.

Kim, M Madden, M and Warner, T 2008, "Estimation of optimal image object size for the segmentation of forest stands with Multi-spectral IKONOS imagery". In: T Blaschke, S. Lang, and G.J. Hay, eds. *Object-based image analysis concepts for knowledge driven remote sensing applications*. Berlin, 2008, pp 291-307, Springer, Berlin, 2008.

Lucieer, A 2004, "Uncertainties in Segmentation and Their Visualisation". PhD Thesis, Utrecht University, ITC Dissertation 113, Enschede, 174 p. http://www.itc.nl/library/papers_2004/phd/lucieer.pdf (accessed 05 August 2013).

Martha, TR Kerle, N Van Westen,CJ Jetten,V and Kumar, KV 2011, "Segment Organization and Data-Driven Shareholding for Knowledge-based landslide Detection by Object-based Image Analysis", *IEEE Transactions of Geo-science and Remote Sensing*, Vol.49, No.12, December 2011.

Marpu, PR Neubert, M Herold, H and Niemeyer, I (2010)," Enhanced evaluation of image segmentation results", *Journal of Spatial Science*, Vol. 55, No. 1, pp. 55–68.

Meinel, G and Neubert, M 2004, "A comparison of segmentation pro- grams for high resolution remote sensing data", *International Archives of Photogrammetry and Remote Sensing*, XXXV, 1097-1105.

Neubert, M and Herold, H 2008," Assessment of remote sensing image segmentation Quality", *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXVIII-4/C1 on CD

Pekkarinen, A 2002, "A method for the segmentation of very high spatial resolution images of forested landscapes", *International Journal of Remote Sensing*, Vol. 23 No. 14, pp.2817-2836.

Woodcock, CE and Strahler, AH 1987, "The factor of scale in Remote Sensing", *Remote sensing of Environment*, Vol. 21 No. 3, pp.311-332.