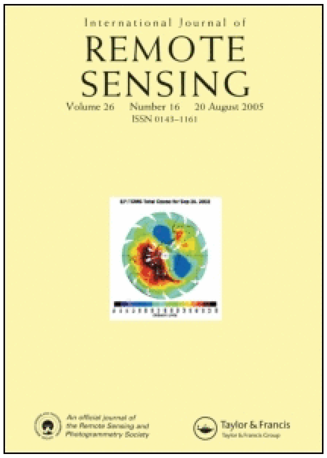

International Journal of Remote Sensing
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713722504

Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition

J. Tian [a]; D. -M. Chen [a]
[a] Department of Geography, Queen's University, Kingston, Ontario K7L 3N6, Canada

Online Publication Date: 01 January 2007
To cite this Article: Tian, J. and Chen, D. -M. (2007) 'Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition', International Journal of Remote Sensing, 28:20, 4625 - 4644
To link to this article: DOI: 10.1080/01431160701241746
URL: http://dx.doi.org/10.1080/01431160701241746

International Journal of Remote Sensing
Vol. 28, No. 20, 20 October 2007, 4625–4644

Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition

J. TIAN* and D.-M. CHEN
Department of Geography, Queen's University, Kingston, Ontario K7L 3N6, Canada

(*Received 16 May 2006; in final form 13 September 2006*)

Multi-resolution segmentation, as one of the most popular approaches in object-oriented image segmentation, has been greatly enabled by the advent of the commercial software, eCognition. However, the application of multi-resolution segmentation still poses problems, especially in its operational aspects. This paper addresses the issue of optimization of the algorithm-associated parameters in multi-resolution segmentation. A framework starting with the definition of meaningful objects is proposed to find optimal segmentations for a given feature type. The proposed framework was tested to segment three exemplary artificial feature types (sports fields, roads, and residential buildings) in IKONOS multi-spectral images, based on a sampling scheme of all the parameters required by the algorithm. Results show that the feature-type-oriented segmentation evaluation provides an insight to the decision-making process in choosing appropriate parameters towards a high-quality segmentation. By adopting these feature-type-based optimal parameters, multi-resolution segmentation is able to produce objects of desired form to represent artificial features.

1. Introduction

Remotely sensed imagery is not composed of semantically discrete real-world objects. Instead, it contains a grid of square pixels that only exhibit simple topological adjacency (Hay *et al.* 2003). Pixel-based image-processing algorithms are argued as being restricted because they only apply on pixels or square areas (Benz *et al.* 2004). Alternatively, segmenting an image into meaningful objects makes it possible to create more informative attributes such as shape, texture, and contextual information. Segmentation is widely adopted as an essential process for most subsequent image-analysis tasks (Haris *et al.* 1998). It divides an image into regions that are expected to correspond to structural units in the scene (Russ 1999). In remote sensing, image segmentation is desired to provide meaningful object primitives for further feature recognition and thematic classification.

Numerous segmentation algorithms have been developed during the past few years. In the taxonomy of segmentation algorithms, they can be broadly grouped into three categories: point-based (Mardia and Hainsorth 1988), edge-based (Jain 1989, Le Moigne and Tilton 1995), and region-based techniques (Chen *et al.* 1991). Point-based or pixel-based segmentation is conceptually the simplest approach (Jèhne 2005). It separates the pixels of an image into different segments by thresholding. Edge-based segmentation approaches track image edges and link them into contours to represent the boundaries of image objects. Region-based

*Corresponding author. Email: 4jt2@qlink.queensu.ca

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/01431160701241746

segmentation mainly includes region merging and split-and-merge approaches. Both detect the regions that meet certain homogeneity criteria. Hybrid approaches are also proposed to improve segmentation performance (Pavlidis and Liow 1990, Shandley *et al.* 1996, Fan *et al.* 2001). In addition, some other approaches based on stochastic theory are also found in the literature (Kraaikamp *et al.* 2001, Wang and Wang 2004); they assume a true image to be a realization of a Markov or Gibbs random field with a distribution that captures the spatial context of the scene (Haris *et al.* 1998).

Multi-resolution segmentation belongs to the category of region-based techniques. The algorithm is characterized as a bottom-up region-merging process starting with one-pixel objects (Baatz *et al.* 2004). Smaller image objects are subsequently merged into larger ones, forming segmentations with objects in different scales. Multi-resolution segmentation has been adopted by a great number of researchers (Darwish *et al.* 2003, Giada *et al.* 2003, Shackelford and Davis 2003, Wang *et al.* 2004, Al-Khudhairy *et al.* 2005, Kressler *et al.* 2005, Wei *et al.* 2005). However, the application of the algorithm still poses several problems. One major characteristic of the algorithm is the ability to segment an image into objects of similar scales at the same level and objects of different scales across levels (the algorithm and associated parameters will be discussed in detail in the following section). Yet, remotely sensed images, especially high-resolution images, often include multiple levels of features that are best addressed at different spatial scales (Chen *et al.* 2004). From the perspective of feature recognition and extraction, for instance, a single residential building differs so greatly from a large shopping mall in terms of scale or size that they are unlikely to be depicted simultaneously on just one single level of segmentation. Consequently, it would be wise to extract the object primitives for different types or classes of features from different segmentation levels. This raises an interesting issue of what would be the best segmentation level for producing optimal object primitives for a given feature class. Until now, there has been little effort devoted to solving this problem in the literature. This may be due to the still-vague definition of segmentation quality in the context of remote-sensing analysis.

This paper proposes a framework of finding an optimal multi-resolution segmentation for a given feature type. The core algorithm of multi-resolution segmentation is introduced in the following section. The proposed framework starts with the definition of meaningful objects for each of the feature types. IKONOS multi-spectral images are used as experimental data. Sports fields, roads, and residential buildings serve as the exemplary feature types to demonstrate the principles. A series of experiments is designed, based on a sampling scheme of all the parameters required by the algorithm. The optimal segmentations are then found by the proposed feature-type-based segmentation assessment.

2. Algorithm of multi-resolution segmentation

Multi-resolution segmentation algorithm has been developed and implemented in the commercial software eCognition (Baatz *et al.* 2004). The idea of this algorithm is to minimize the average heterogeneity of image objects weighted by their size during the merging process. The merges are performed in a pairwise manner. For each possible merge of two adjacent objects, the heterogeneity change, also referred to as 'merging cost' in Baatz and Schäpe (2000), is quantitatively evaluated and then compared with a parameter (scale parameter S_p) specified as the threshold. A possible merge is fulfilled if its heterogeneity change is smaller than the given value of the parameter. Hence, a larger value allows for more merges than a smaller value, leading to the production of larger objects.

2.1 *Measurement of heterogeneity change*

In the algorithm, the heterogeneity change caused by a possible merge is evaluated by calculating and mixing the spectral and shape differences between the situations before and after the merge. For any pair of adjacent objects, an overall heterogeneity change for a possible merge is broken down into two contributors, which are linearly combined to calculate a value as the following.

$$\Delta h_{\text{overall}} = \left(1 - W_{\text{shape}}\right) \Delta h_{\text{spectral}} + W_{\text{shape}} \Delta h_{\text{shape}} \tag{1}$$

where W_{shape} is an arbitrary weight of importance ranging from 0 to 1. $\Delta h_{\text{spectral}}$ and Δh_{shape} measure the heterogeneity changes of spectral and shape, respectively. They are defined as follows:

$$\Delta h_{\text{spectral}} = \sum_{i=1}^{N} W_{\text{i}} \left(n_{\text{Merge}} \sigma_{\text{i}}^{\text{Merge}} - \left(n_{\text{Obj1}} \sigma_{\text{i}}^{\text{Obj1}} + n_{\text{Obj2}} \sigma_{\text{i}}^{\text{Obj2}} \right) \right) \tag{2}$$

$$\Delta h_{\text{shape}} = W_{\text{cmpct}} \Delta h_{\text{cmpct}} + \left(1 - W_{\text{cmpct}}\right) \Delta h_{\text{smooth}} \tag{3}$$

N denotes the number of channels of an image to be segmented. W_i represents the weight assigned to the ith channel, and n denotes the number of pixels belonging to the object. In the algorithm, the spectral heterogeneity of an object is measured by the standard deviation of the spectral values of the pixels making up the object. The terms σ_i^{Obj1}, σ_i^{Obj2}, and σ_i^{Merge} are used to represent the respective standard deviations during a pairwise merge (equation (2)).

As can be seen in equation (3), the measure of shape heterogeneity change is further broken down into two sub measures: compactness change Δh_{cmpct} and smoothness change Δh_{smooth}. These two measures are also linearly weighted in the same way as equation (1). In multi-resolution segmentation, the compactness of an object is measured by calculating the deviation from the ideal compact form given by the relation between factual border length l and the root of the object size n in pixels (Benz *et al.* 2004). Meanwhile, smoothness is measured by the ratio of factual border length l over the perimeter b of the bounding box. (The bounding box is defined as the rectangle that is just large enough to contain all object pixels in Jèhne 2005.) The measures Δh_{cmpct} and Δh_{smooth} are therefore defined and obtained in a comparative manner illustrated by equations (4) and (5).

$$\Delta h_{\text{cmpct}} = n_{\text{Merge}} \frac{l_{\text{Merge}}}{\sqrt{n_{\text{Merge}}}} - \left(n_{\text{Obj1}} \frac{l_{\text{Obj1}}}{\sqrt{n_{\text{Obj1}}}} + n_{\text{Obj2}} \frac{l_{\text{Obj2}}}{\sqrt{n_{\text{Obj2}}}} \right) \tag{4}$$

$$\Delta h_{\text{smooth}} = n_{\text{Merge}} \frac{l_{\text{Merge}}}{b_{\text{Merge}}} - \left(n_{\text{Obj1}} \frac{l_{\text{Obj1}}}{b_{\text{Obj1}}} + n_{\text{Obj2}} \frac{l_{\text{Obj2}}}{b_{\text{Obj2}}} \right) \tag{5}$$

2.2 *Decision heuristics of merging*

Equations (1)–(5) govern the local computation and decision-making for all the possible pairwise merges. In order to trigger the merging process and reach a repeatable segmentation result when given a certain set of parameters, some objects are designated as points of departure. Furthermore, the algorithm treats image objects in a distributed order (Baatz and Schäpe 2000). Multi-resolution segmentation allows all image objects to grow simultaneously, so that they will always be of comparable size. For more details, refer to Baatz *et al.* (2004) and Benz *et al.* (2004).

2.3 *Algorithm discussion*

The principles described in §2.1 and 2.2 ensure a steady and unique segmentation result given a certain set of the parameters (W_{shape}, W_{cmpct}, and S_p) to an image. It is noteworthy that most of the common segmentations have difficulties in reproducing the same result for the same region from different subsets. The same image area being part of different subsets will be segmented differently by means of these procedures because they are dependent on the particular feature space of the respective subset (Baatz *et al.* 2004). Multi-resolution segmentation is superior at this point, segmenting the same area in the same or very similar way, no matter how large the entire scene under processing.

Moreover, the algorithm has the capability of considering not only the spectral information but also the shape information for each possible merge. Although the account of shape information to some degree may lead to a better segmentation result, it is hard to determine how much weight the shape information should be given and what kind of shape information should mainly be used. The value choosing for the parameters W_{shape} and W_{cmpct} may strongly affect the form of the resulting objects from a segmented image. Baatz *et al.* (2004) suggest emphasizing the *spectral* information as much as possible while keeping the *shape* information as much as necessary to produce image objects. The reason is because the use of shape information works at the cost of spectral homogeneity; using too much shape information can therefore reduce the quality of the segmentation results. This suggestion is recognized as a 'rule of thumb' short of quantitative explanation or evidence. There is little insight as to how large or small the parameters should be in order to obtain the optimal results to purpose, and how sensitive the results are to the parameters' change. These questions will be answered, even if not ultimately, in this research.

3. Methodology

In order to be valuable to the following classification or feature extraction, an image should ideally be segmented into so-called meaningful objects (Baatz *et al.* 2004). An immediate problem that researchers need to confront is how to define or assess an object's meaningfulness. The definition of meaningful objects therefore constitutes the first step towards the evaluation of a segmentation in this research. The second step is to perform a series of experimental segmentations based on a sampling scheme of the algorithm-required parameters. Third, a measure, termed G_{s}, is proposed to quantify the goodness of the segmentations. By comparing to the desired feature shape (digitized polygon), the segmentation results can be quantitatively evaluated by G_{s}. The optimal parameters can thus be found out and applied to other test images.

3.1 *Define meaningful objects*

Intuitively, an object should be defined as meaningful if it precisely depicts its background feature in an appropriate shape and size. This shape should satisfy the human eye and can generally be understood as an instance of the conceptual recognition of the feature class. Furthermore, the conceptual recognition of different feature classes is rooted in their different regularities in terms of shape and structure. Hence, it seems unwise to define an object's meaningfulness in general; rather, a feature-class respective definition might be more suitable. In other words, the same object may be meaningless to one feature class, but meaningful to another. As a comprehensive definition for all the feature types is a demanding task, it was decided that only the classes of sports fields, roads, and residential buildings, representing features with different complexity levels and shapes, are to be included to demonstrate the principles.

3.1.1 Sports fields. Among the three feature classes, sports fields are thought of as having the least complexity in both shape and composition. Although normally covering a fairly large area, the central fields surrounded by tracks are spatially intensive, usually found in an elliptical shape. The object representing such a central field should be defined as meaningful when it has a relatively high elliptic fit and is large enough. The elliptic fit can be measured based on the creation of an ellipse with the same area and same ratio of length/width as the considered object. The fitness can then be reflected by the ratio of the area of the object over the area of the object portion inside the ellipse (Baatz *et al.* 2004). The ratio ranges from 0 to 1: the higher the ratio, the better the elliptic fit. In figure 1, the central field is represented by a meaningful object in the right graph but four non-meaningful objects in the middle.

3.1.2 Road. Roads have a very strong regularity. In remote-sensing images, roads, intended to form networks in urban areas, are generally recognized as linear, connective, and extensive. Regardless of their structural complexity, roads are essentially recognized by their linearity. An object should therefore be defined as a meaningful road when it is rather linear. It seems fair to say that this definition actually abstracts road segments rather than the overly broad concept of road. In practice, however, it is not so straightforward to quantitatively capture an object's linearity.

Road segments can be reasonably categorized into two groups of conceptual models: the objects of connection and the objects of intersection. The typical shapes of the former and the latter are shown by (*a*), (*b*) and (*c*), (*d*), respectively, in figure 2.

Figure 1. Comparison of a sports field shown on the IKONOS image subset (left), and the meaningless (middle) and meaningful (right) objects of the central field. Available in colour online.

Figure 2. Four typical shapes of road objects. The upper two represent typical objects of connection; the lower two represent typical objects of intersection. The thicker inner lines represent the main lines. Available in colour online.

It is found that, for a given object, depending on type, its linearity can be represented by one of the following three object features (measures):

1. *Length/width of object*: in pattern recognition, the length/width of a pixel-based object can be measured by the fraction of the eigenvalues of the object-derived covariance matrix, with the larger eigenvalue being the numerator (Baatz *et al.* 2004). As segmentations produce raster polygons rather than polygons in free shape, the length/width of an object from segmentation can be measured using the same method as well. However, this measure is not able to capture the linearity of different objects. Only the linearity of a fairly straight object along a grid, like figure 2(*a*), can be well reflected by this object feature. In contrast, the linearity of an actually linear object cannot be reflected by this object feature when the object is curved to some degree.
2. *Length/width of main line*: the characteristics considering the objects' main line are superior in recognizing curved linear objects in the shape of figure 2(*b*) or likewise. The main line of an object can be determined by its skeleton. In operation, the creation of the skeleton can be implemented based on a Delaunay triangulation (Worboys and Duckham 2004) of an object's shape polygon. The skeleton is then obtained by identifying the mid points of the triangles and connecting them. Meanwhile, all the points can be classified into branch point, connection point, and end-point, according to the number of

neighbours of their corresponding triangle. Thus, the main line can be reasonably represented by the longest possible connection of branch-points. The other connected lines can then be ordered in the same manner as ordering the streams in a basin/watershed. A typical object skeleton therefore consists of one main line and some branches of first order and up. For more details, refer to Baatz *et al.* (2004). The length of an object's main line is simply the sum of the distances between the nodes. To calculate the length/width of the main line, the width of an object can be estimated by the average height of all triangles crossed by the main line. This object feature captures the linearity of a linear object of free direction.

3. *Length/width of main line plus long branches of first order*: although recognized as typical road intersections, the objects shown in figure 2(*c*) and figure 2(*d*) are not simply linear. More rigorously, they should be regarded as structured objects consisting of linear object components. The main line may not solely be able to reflect their linear nature. A new feature is therefore developed, taking into account not only the main line but also the long branches of first order. For a given object, this feature needs to measure the length of the main line L_{Main}, the total length of all the long branches of first order $L^1{}_{\mathrm{Branch}}$, and the object area A. The feature value thus equals $(L_{\mathrm{Main}}+L^1{}_{\mathrm{Branch}})^2/A$. There is one long branch and two long braches for the highlighted objects in figure 2(*c*) and figure 2(*d*), respectively.

For a given object, the examination of all of the above three object features helps determine more confidently whether it is linear or not. An object should be defined as a meaningful road with strong confidence if it has a relatively large value for at least one of the three associated features.

3.1.3 Residential buildings with simple roof structures. From the point view of remote sensing, most often a building is actually visible as the building's roof. In general, building roofs can be very complex and comparatively have the least regularity in terms of shape, size, and structure. Gruen and Dan (1997) classify the models of building roofs according to their number of ridge points. Only buildings with simple roof structures are covered in this research. More particularly, the majority of the buildings investigated have a rectangular footprint. The roofs are flat or have one or two ridge points. Due to the diversity of building roofs, the definition of a building's meaningfulness is purely based on the comparison of the desired model of the building and the actual object primitive prepared by the segmentation. It is decided here that an object taking the place of a building is defined as meaningful only if the object has a high percentage of mutual overlap with the desired model (e.g. digitized polygon).

3.2 *Assessment of segmentations*

A multi-resolution segmentation with a certain set of parameters may be good at providing meaningful objects for one feature class but poor for another. Regarding the definition of meaningful objects, the segmentations should be wisely evaluated according to the feature class of interest. Our strategy is to compare the collection of the relevant and meaningful objects to a feature class of interest with its predetermined polygons of reference. An object is regarded as relevant to a feature type if its majority area falls in the corresponding reference polygons; on the other hand, the meaningfulness of the object is assessed by the principles proposed in §3.1. For a feature class of interest on a given image, the reference polygons can be obtained from other resources or by digitization.

The method of assessment based on an overlay analysis is illustrated by figure 3. The black rectangle underneath represents one reference polygon of the feature class investigated. The forward white rectangles represent the relevant and meaningful objects to the feature class. The grey portions labelled A_{overlap} are the overlapping areas, while the portions labelled A_{diff} are the areas mismatching. Normally, there should be only one meaningful object supposed to take the place of the background feature when it is intensive (e.g. buildings). If a feature is spatially extensive (e.g. roads), it is likely to be covered by multiple relevant and meaningful objects.

A measure was developed to quantify the goodness of the segmentation. The measure, termed G_s, takes into account both the overlapping areas and the mismatching areas. A perfect collection of the meaningful objects will yield a G_s value of 1 from equation (8). At the other extreme, if there are no meaningful objects overlapping with the reference polygons, G_s will yield 0. In all other cases, the value will be in between.

$$A_{\text{overlap}} = \sum A_{\text{i}}^{\text{overlap}} \tag{6}$$

$$A_{\text{diff}} = \sum A_{\text{i}}^{\text{diff}} \tag{7}$$

$$G_{\text{s}} = \frac{A_{\text{overlap}}}{A_{\text{refer}} \cdot e^{A_{\text{diff}}/A_{\text{refer}}}} \tag{8}$$

Overlay analysis plays a key role in obtaining the overlapping area and the difference area required by calculating G_s. Given two layers of vector polygons, one the working layer and the other the identity layer, the analysis first computes their geometric intersection. The intersected polygons on the working layer will then be split into more portions. In the output, the portions common to the two layers will become new polygons and inherit the attributes including their 'area'. Particularly in this research, the vector from the segmentation and the vector of reference serve as the working layer and the identity layer, respectively. When a new polygon of the overlapping portion is common to a relevant and meaningful object and a reference polygon, the area of the new polygon is counted as an overlapping area A_{overlap}. Simultaneously, the difference between the inherited area of the original meaningful polygon and the area of the new polygon is reasonably counted as a difference area

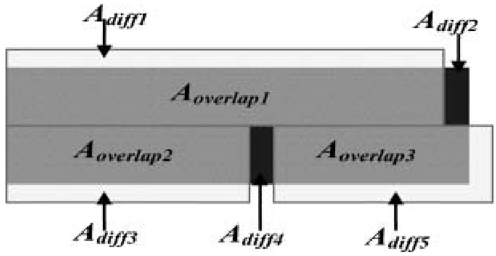

Figure 3. Sketch for the overlap analysis.

A_{diff}. It should be pointed out that A_{diff} also includes the area of the portions that are part of a reference polygon but not covered by any meaningful object from the segmentation result.

4. Experiment design

4.1 *Experimental-data description*

The IKONOS-2 satellite has become an important resource providing high-resolution space images to the remote-sensing community. Two IKONOS satellite images were employed in this research to perform experiments. One has a spatial resolution of 4 m, covering the western part of Kingston, Ontario, Canada. Six sites were subset from this image, with three representing road areas and the other three representing areas of sports field. As the 4-m resolution does not seem fine enough to detect either residential or commercial buildings, a Pan-sharpened multispectral IKONOS image with 1-m resolution was also employed, covering part of San Diego, California. On this image, individual buildings could be visually recognized. Two sites of residential buildings were subset from the image. For each feature type, one site is used for testing the methodology, and the remaining sites are used for validating the findings from the test site.

Both IKONOS images have four spectral bands: blue, green, red, and near infrared. The spectral values in each band are in 11 bits.

4.2 *Segmentations based on a series of parameter settings*

A set of parameters need to be specified in the algorithm of multi-resolution segmentation. The choice of plane four-neighbourhood or diagonal eight-neighbourhood determines the definition of neighbourhood for most merging-based segmentations. Baatz *et al.* (2004) suggest that the segmentation be performed in the manner of the plane four-neighbourhood on a high-resolution image. With respect to this, there are still four parameters open to specification: W_{shape}, W_{cmpct}, W_{smooth} (form-related parameters), and S_p (size-related scale parameter). As W_{cmpct} and W_{smooth} are trade-off parameters totalling 1, the specification of only one of the two is needed. Thus, an image will be segmented differently based on different settings of three independent parameters. If W_{shape} is set to be zero, this means that there will be no shape information considered when segmenting. The sub-parameters of W_{shape} (refer to equation (3)), W_{cmpct}, and W_{smooth} can consequently be neglected in that case because they will not impact the segmentation result. To the contrary, if W_{shape} is set to be non-zero, this means that shape information will be taken into account. The volume of W_{shape} indicates how much the segmentation will emphasize the shape. Four values were sampled to explore the impact of W_{shape} on the segmentation: 0.25, 0.5, 0.75, and 0.9. Notice that the largest value is not 1 because no spectral information will be considered if that is the case. It is understandable that the segmentation will not make sense if no spectral information is considered at all during the course of segmentation. Each non-zero setting of W_{shape} requires a sub-setting of W_{shape} and W_{cmpct}. Five values (0, 0.25, 0.5, 0.75, and 1) were chosen for W_{cmpct} covering the range of 0–1 with an interval of 0.25.

In total, there are 21 different settings of form-related parameters. Among them, only Setting 0 does not take any shape information into account. All the other 20 settings are summarized in table 1, where the columns and the rows show the

Table 1. Experimented parameter settings of W_{shape} and W_{cmpct}.

Shape parameter	Parameter of compactness				
	0.00	0.25	0.50	0.75	1.00
0.25	Setting 1	Setting 2	Setting 3	Setting 4	Setting 5
0.50	Setting 6	Setting 7	Setting 8	Setting 9	Setting 10
0.75	Setting 11	Setting 12	Setting 13	Setting 14	Setting 15
0.90	Setting 16	Setting 17	Setting 18	Setting 19	Setting 20

There is one more setting with a shape parameter of 0.

experimental values for W_{shape} and W_{cmpct}, respectively. All 21 settings were applied to the subset images. Under each setting, the images were segmented on 70 scale levels, with the corresponding scale parameters of S_p ranging from 0 to 145. Thus, a framework was built up to facilitate tracking the quality of the segmentations based on the different parameter combinations.

All the segmentations were performed in eCognition. This software allows for editing protocols that enable a suite of processes to be completed in just a few batches. The multi-resolution segmentation results can be directly exported as vector data in a GIS-compatible format, such as shape file. This process could be made much more efficient by editing and using protocols. It is found that eCognition has the disadvantage of exporting image polygons without any spatial reference information attached. The georegistration of a shape file exported from eCogntion has to be completed in a GIS package, e.g. ArcGIS. A customized tool was developed in the ArcGIS environment to programmatically assess the segmentations' quality.

5. Result analysis and discussion

5.1 *Sports field-oriented segmentation*

Experiments were performed on one of the three areas of sports field subset from the original 4-m IKONOS image. Figure 4 exhibits the impact of the settings on the segmentation quality for the central field within the tracks. It can be seen that most of the settings have the capability of producing fairly good results. The best segmentation results achieved from the different settings are on a similar quality level (with G_s about 0.9 or above). The segmentation quality does not continuously change as the scale parameter increases. When the scale parameter is small, the corresponding segmentation quality appears to be 0, indicating that there are no meaningful objects to represent the central field at all yet. Most of the changing lines climb in one or two cascades up to their maximum of about 0.9 in the vertical axis. The maxima of all of the changing lines appear to be flat. After staying at the maximum across some scale parameters, the changing lines then drop sharply down to 0. From graph (*a*) to (*d*) in figure 4, the importance of the shape is increasingly emphasized. It was noticed that the segmentations under a setting of higher W_{shape} reach their best at comparatively smaller scale parameters. Moreover, the tops of the lines are consistently flat but shorter from (*a*) to (*d*), indicating that the optimal segmentations occur within a gradually narrower range of scale parameter.

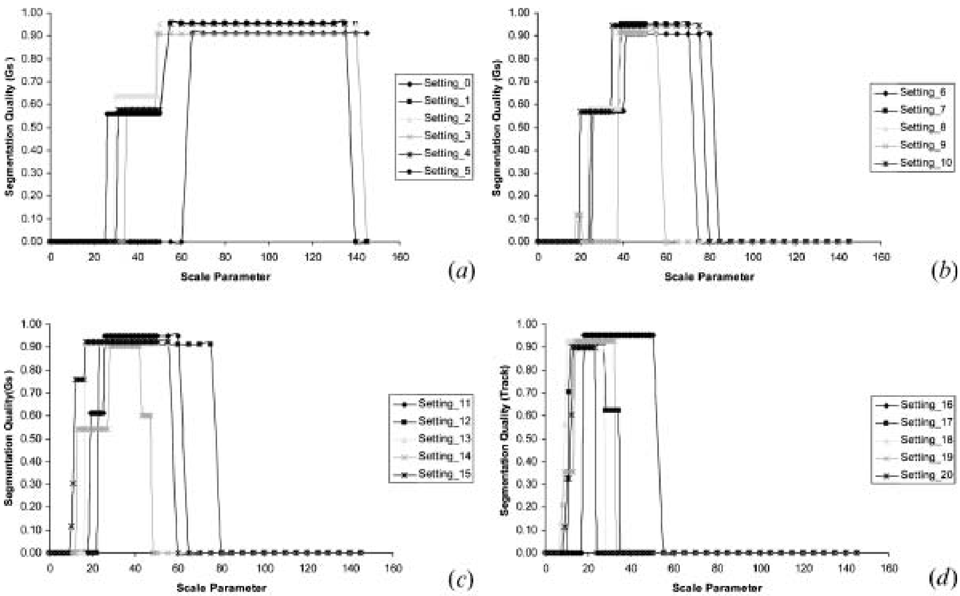

Figure 4. Quality of sports field-oriented segmentations against scale parameter.

Table 2. Goodness of the best segmentations for the sports field and their associated scale parameters.

	Parameter of compactness									
	0.00		0.25		0.50		0.75		1.00	
Shape parameter	G_s	S_P	G_s	S_P	G_s	S_P	G_s	S_P	G_s	S_P
0.25	0.95	55	0.95	50	0.91	49	0.95	55	0.96	55
0.50	0.91	41	0.95	39	0.91	39	0.91	38	0.95	35
0.75	0.95	26	0.91	23	0.92	18	0.90	28	0.92	17
0.90	0.95	18	0.92	12	0.93	10	0.93	14	0.90	13

Table 2 summarizes the 'goodness' values of the best segmentations under all the settings. The corresponding scale parameters are also included in the table. It seems fair to say that, for sports fields, the segmentations with a higher W_{shape} are more efficient in achieving a high-quality result. The reason is that the segmentations associated with a higher W_{shape} produce comparable results with smaller scale parameters, which allow for fewer merges. Hence, based on consideration of both segmentation quality and computing efficiency, the parameter setting 16 (W_{shape}=0.9 and W_{cmpct}=0) is regarded as the ideal one to delineate the central field within the tracks. The same setting was applied to the other two subsets of sports field at the same scale parameter of 18. The segmentation results (figure 5(*d*) and 5(*f*)) are very good, thus supporting the strength of the setting and the chosen scale parameter.

5.2 *Road-oriented segmentations*

Similarly, a series of experiments was performed on one of the three road area (figure 6(*a*)) subsets from the 4-m IKONOS image. Figure 7 summarizes the impact of the different settings on the segmentation quality for the roads. From graph (*a*) to

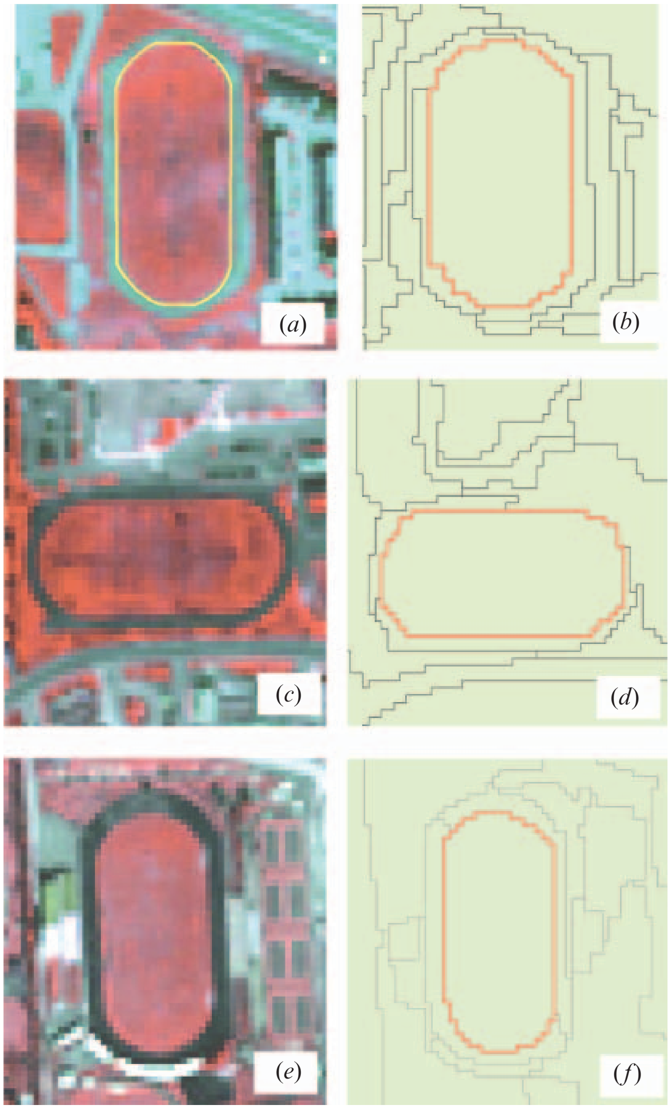

Figure 5. Segmentation results of the three subsets of sports fields with W_{shape} of 0.9 and W_{cmpct} of 0 at a S_p of 18. The image sizes are (*a*) 51 × 52, (*b*) 56 × 52, (*c*) 53 × 57. Image (*a*) is overlaid with its reference polygon in yellow. The meaningful objects segmented for the central field are highlighted in red.

(*d*) in the figure, the associated importance of shape is increasingly emphasized. Regardless of the change of settings, all the changing lines exhibit a unimodal distribution with the lower ends and most of the upper ends touching the *x*-axis. This indicates that scale parameter with very small or very large values are not effective in producing meaningful road objects. Very small scale parameters allow only the creation of the objects in size of pixel level; apparently, these objects are not representative of the roads in terms of size or shape. Contrarily, at the other extreme, the large scale parameters are intended to form such large objects that often include features of different types into one object due to relatively large tolerance of heterogeneity changes. The existence of the modes is somewhat in

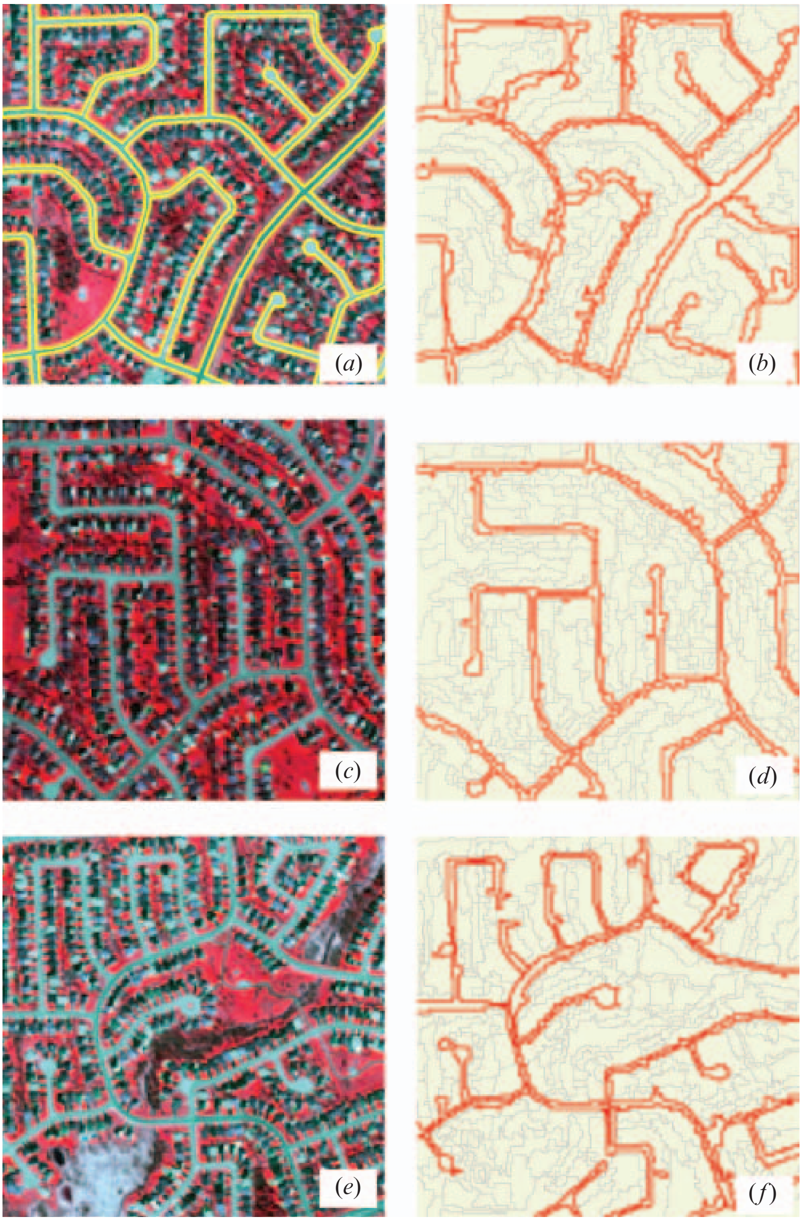

Figure 6. Original IKONOS subsets of roads (left) and the segmentation results (right) at S_p of 23, W_{shape} of 0.75, and W_{cmpct} of 0.25. The image sizes are (*a*) 151×150, (*b*)148×138, (*c*) 171×175. The image (*a*) is overlaid with the corresponding reference polygons. The meaningful objects segmented for the roads are highlighted in red.

parallel with the hypothesis that one feature type is best addressed at a certain scale or within a certain scale range.

The six groups of segmentations shown in figure 7(*a*) have a relatively weak emphasis (W_{shape}=0 or 0.25) on the importance of shape. Notably, the 'belt' of the high–low range becomes narrow at scale parameters of around 30, forming a 'neck' area. Away from the area, the belt spreads out, reflecting a stronger impact of

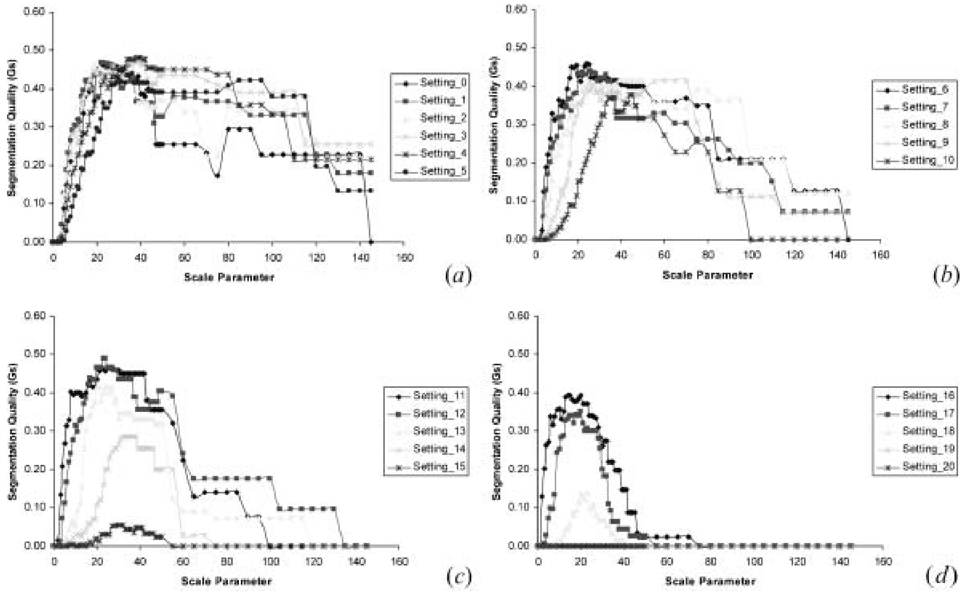

Figure 7. Quality of the road-oriented segmentations against the scale parameter.

W_{cmpct} on the segmentation quality. When the importance of shape and spectral information is emphasized on a basis of half and half, as in figure 7(*b*), the segmentation with a W_{cmpct} of 0 generates the best result overall. It can be seen that a 'neck' area still exists, although it seems less significant. The 'neck' areas are recognized as being informative. They imply a steady range of scale parameters, where the segmentation quality is not very sensitive to the different settings.

When the importance of shape is weighted over half, as in figure 7(*c*) and (*d*), there no longer exists a 'neck' area. However, a trend appears clearer that the segmentation quality for road decreases as the associated W_{cmpct} increases. Especially in graph (*d*), the trend is unexceptionally followed. Also, Setting 19 and Setting 20 have their changing lines lying on the *x*-axis, implying a total failure of producing any meaningful object for the feature class of road.

From table 3, the segmentations with W_{shape} of 0.25 produce comparatively the best results with the highest average of the goodness values. The segmentations with W_{shape} of 0.5 produce slightly worse results. The segmentations with W_{shape} of 0.75

Table 3. Goodness of the best segmentations for the roads and their associated scale parameters.

Shape parameter	Parameter of compactness 0.00 G_s	0.00 S_P	0.25 G_s	0.25 S_P	0.50 G_s	0.50 S_P	0.75 G_s	0.75 S_P	1.00 G_s	1.00 S_P
0.25	0.47	22	0.47	19	0.47	37	0.48	38	0.45	32
0.50	0.46	24	0.44	25	0.42	39	0.42	60	0.38	44
0.75	0.47	25	0.49	23	0.42	20	0.29	32	0.05	29
0.90	0.39	14	0.35	20	0.14	23	0.000	N/A	0.000	N/A

produce relatively good results when W_{cmpct} is low but very poor results when W_{cmpct} is high. In contrast, the segmentations with W_{shape} of 0.9 result in the worst. All in all, the best segmentation occurred at a scale parameter of 23 with W_{shape} of 0.75 and W_{cmpct} of 0.25. The same setting was applied to the other two image subsets of roads (figure 6(*c*) and 6(*e*)), and their segmentation results are shown by figure 6(*d*) and 6(*f*), respectively.

5.3 *Building-oriented segmentations*

Buildings differ very much from one another in terms of roof structure and colour. The building roofs investigated in this research have a rectangle-like footprint, and most of them are homogeneous in colour. As can be seen in figure 8, all the changing lines are basically unimodal. This indicates that the segmentation under any of the settings increases its quality for building when the scale parameter increases, until the value corresponding to the curve top is reached. From there to the right, the segmentation quality decreases eventually to 0 when the corresponding scale parameter is quite large. Again, from (*a*) to (*d*) in figure 8, the shape is increasingly emphasized when segmenting. Overall, it is noticed that the best segmentation quality associated decreases from (*a*) to (*d*).

Table 4 summarizes the best segmentations associated with the designated parameter settings for the buildings. The best segmentation occurred when no shape information was considered, and the scale parameter was 43. Examination of the segmentation result shows that most of the buildings are well represented by polygons in shapes close to what are desired (see the left-hand side part of figure 9). For building-oriented segmentation, shape information seems useless. However, closer and more specific examination reveals that one building was not properly

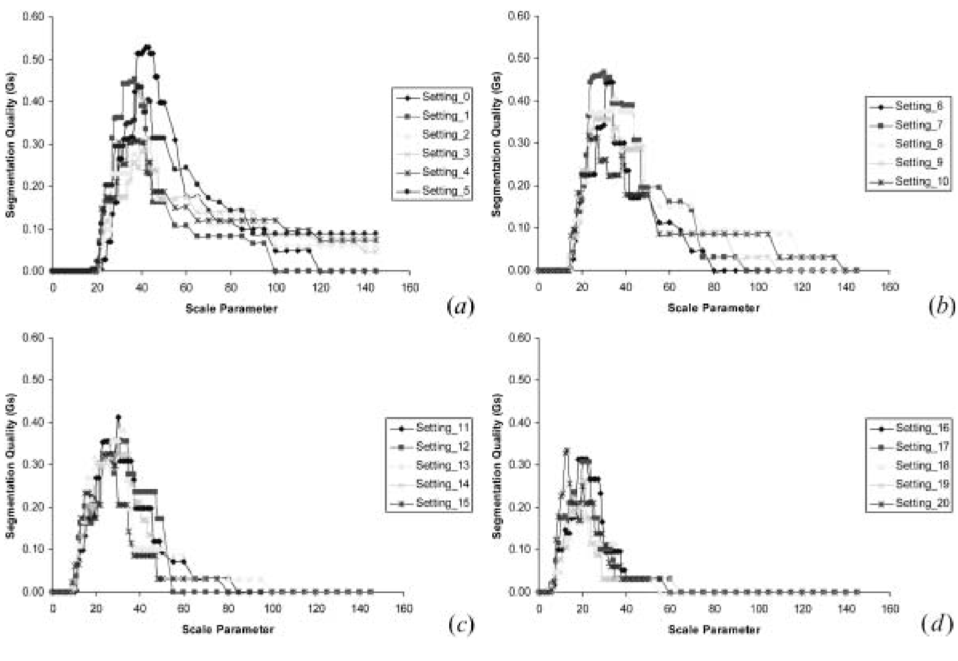

Figure 8. Quality of the building-oriented segmentations against the scale parameter.

Table 4. Goodness of the best segmentations for the buildings and their associated scale parameters.

Shape parameter	Parameter of compactness 0.00		0.25		0.50		0.75		1.00	
	G_s	S_P	G_s	S_P	G_s	S_P	G_s	S_P	G_s	S_P
0.25	0.45	37	0.45	40	0.31	43	0.31	37	0.43	38
0.50	0.44	32	0.47	30	0.41	35	0.36	30	0.32	23
0.75	0.41	30	0.36	30	0.38	31	0.33	29	0.33	23
0.90	0.31	18	0.31	20	0.22	17	0.26	20	0.36	13

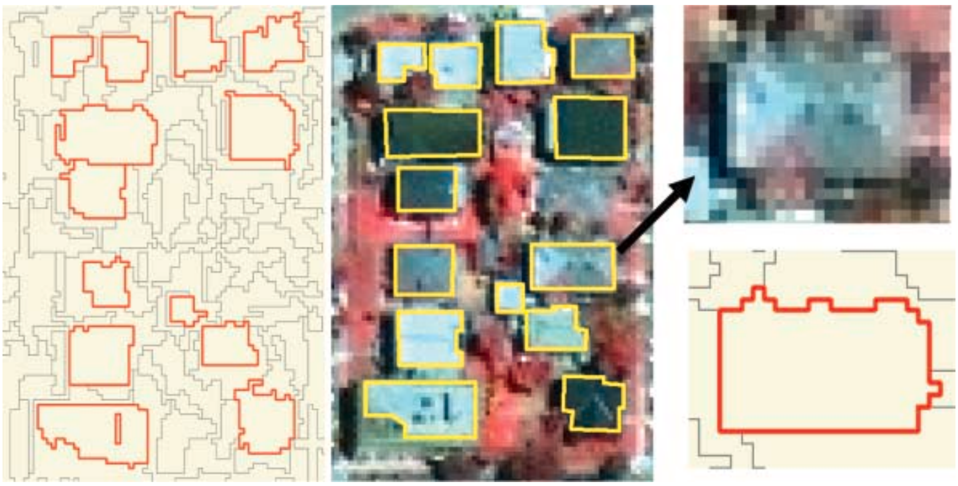

Figure 9. Building-oriented segmentation. The left part shows the segmentation result of the middle image (71 × 106) at S_p of 43 with no shape information considered; the lower right part shows the segmentation result for the pointed building at S_p of 20 with W_{shape} of 0.9 and W_{cmpct} of 0. Available in colour online.

represented by any of the polygons from the segmentation. The building is recognized as heterogeneous in colour and not having strong contrast with the surrounding areas. Although not so clear, the building roof actually has two ridges, so when there is sunshine casting from an angle, the different portions of the roof will be shaded differently from light to dark. It was found that this building could be better represented by taking shape information into the segmentation (see the right-hand side part of figure 9).

It is commonly agreed that building roofs vary greatly in colour, shape, and structure. In other words, building roofs have a much lower regularity compared with sports fields and roads. Hence, the optimal parameters found for building-oriented segmentation in one test image may not hold in another that has quite different building roofs in size or structure. For example, figure 10 shows the segmentation result of the other subset from the 1-m Pan-sharpened multispectral IKONOS image when the scale parameter is specified as 43 and the shape factor is set to be 0 (the same setting as that in figure 9). Only a limited number of meaningful building objects are provided.

Figure 10. Building-oriented segmentation of the test building image subset (75×70) by applying the suggested optimal parameter used in figure 9 (left). Available in colour online.

6. Concluding remarks

The definition of meaningful objects plays a crucial role in building the framework of finding an optimal segmentation for a given type of object. The meaningfulness of an object to three artificial feature types has been accordingly defined in this study. The segmentation quality can thus be evaluated by comparing the meaningful and relevant objects segmented with the corresponding reference polygons. An assessment method has also been proposed, and a series of experiments have been demonstrated.

Artificial objects include not only intensive features (e.g. buildings) but also extensive features (e.g. roads). An intensive feature should ideally be represented by only one single meaningful object taking the place of the feature, whereas an extensive feature will likely be represented by multiple objects meaningful to the feature class. Among the three types of objects investigated, sports fields have comparatively the strongest regularity and uniqueness in terms of shape. They are usually recognized by their elliptical shapes and their great colour contrast with the surrounding area. It was found that a high-quality segmentation result for sports field is not so dependent on the setting of the form-related parameters. No matter how the form-related parameters change, the multi-resolution segmentation algorithm seems capable of producing comparable results. Yet, the best segmentation with different form-related parameters may occur with different scale parameters. While ensuring high quality, the ideal segmentation should also adopt a relatively small scale parameter as it will allow fewer merges and save computing time. The experiments suggest that a W_{shape} of 0.9 and a W_{cmpct} of 0 may be somewhat optimal to segment a 4-m IKONOS multi-spectral image for sports fields.

Compared with sports fields, roads have less, but still strong, regularity. Roads are characterized by their linear and networking nature, making the conceptual models of roads less straightforward to build. Three object features were seen to be valuable in defining meaningful road objects. For the road-oriented segmentations, the importance of shape (W_{shape}) should not be emphasized too strongly. Small or medium values (less than or equal to 0.75) of W_{shape} were found to be suitable to produce fairly good segmentation results for road. Meanwhile, the importance of compactness (W_{cmpct}) should be kept small in most of the cases. The best segmentation occurred at the scale parameter of 23 with W_{shape} equal to 0.75 and W_{cmpct} equal to 0.25. The success of the segmentations of other road image subsets using the same set of parameters supports its robustness.

Buildings are much more complex. Two blocks of simple residential buildings were examined in this research. The experiments imply that a building-oriented segmentation should mainly rely on the spectral information contained in an IKONOS image. However, for those buildings with a roof of heterogeneous colour, the importance of shape should be fully emphasized. The determination of an optimal scale parameter is subject to the size of the buildings of interest. The optimal parameters found for one group of buildings may be of limited value for another with a very different size or structure. The multi-resolution segmentation for buildings should be performed, based on a close examination of the buildings' sizes, roof structures, and colour variances.

In summary, the trend in the segmentation quality changes helps researchers narrow down their range of the parameters' specification. Although the robustness of the suggested optimal parameters needs more examination by a significant number of tests, their usefulness has been preliminarily demonstrated. The proposed framework of finding optimal segmentation to a feature class of interest may practically aid in artificial feature recognition from high-resolution satellite images. The relationship between the image resolution and the optimal parameters, especially the scale parameter, should be further examined in the future work.

Acknowledgements

We would like to thank the funding support from Natural Sciences and Engineering Research Council of Canada. The IKONOS image of Kingston was provided by the Department of Geography, Queen's University. The IKONOS image of San Diego was downloaded from the website of Space Imaging Corporate.

References

Al-Khudhairy, D.H.A., Caravaggi, I. and Giada, S., 2005, Structural damage assessments from Ikonos data using change detection, object-oriented segmentation, and classification techniques. *Photogrammetric Engineering and Remote Sensing*, **71**, pp. 825–838.

Baatz, M., Benz, U., Dehghani, S., Heynen, M., Höltje, A., Hofmann, P., Lingenfelder, I., Mimler, M., Sohlbach, M., Weber, M. and Willhauck, G., 2004, *eCognition Professional: User Guide 4* (Munich: Definiens-Imaging).

BAATZ, M. and SCHÄPE, A., 2000, Multiresolution segmentation—an optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informationsverarbeitung XII, Beiträge zum AGIT-Symposium Salzburg 2000* (Karlsruhe: Herbert Wichmann Verlag), pp. 12–23.

BENZ, U., HOFMANN, P., WILLHAUCK, G., LINGENFELDER, I. and HEYNEN, M., 2004, Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing*, **58**, pp. 239–258.

CHEN, D., STOW, D. and GONG, P., 2004, Examining the effect of spatial resolution on classification accuracy: an urban environmental case. *International Journal of Remote Sensing*, **25**, pp. 2177–2192.

CHEN, S., LIN, W. and CHEN, C., 1991, Split-and-merge image segmentation based on localized feature analysis and statistical tests. *Graphical Models and Image Processing*, **53**, pp. 457–475.

DARWISH, A., LEUKERT, K. and REINHARDT, W., 2003, Urban land-cover classification: an object based perspective. In *2nd GRSS/ISPRS Joint Workshop: Remote Sensing and Data Fusion over Urban Areas*, 22–23 May, Berlin, pp. 278–282.

FAN, J., YAU, D.K.Y., ELMAGARMID, A.K. and AREF, W.G., 2001, Automatic image segmentation by integrating color-edge extraction and seeded region growing. *IEEE Transactions on Image Processing*, **10**, pp. 1454–1466.

GIADA, S., DE GROEVE, T. and EHRLICH, D., 2003, Information extraction from very high resolution satellite imagery over Lukole refugee camp, Tanzania. *International Journal of Remote Sensing*, **24**, pp. 4251–4266.

GRUEN, A. and DAN, H., 1997, TOBAGO-A topology builder for the automated generation of building models, automatic extraction of man-made objects from aerial and space images (11). Birkhäuser, Basel, pp. 149–160.

HARIS, K., EFSTRATIADIS, S., MAGLAVERAS, N. and KATSAGGELOS, A., 1998, Hybrid image segmentation using watersheds and fast region merging. *IEEE Transactions on Image Processing*, **7**, pp. 1684–1699.

HAY, J., BLASCHKE, T., MARCEAU, J. and BOUCHARD, A., 2003, A comparison of three image-object methods for the multiscale analysis of landscape structure. *ISPRS Journal of Photogrammetry & Remote Sensing*, **57**, pp. 327–345.

JAIN, A., 1989, *Fundamentals of Digital Image Processing* (Englewood Cliffs, NJ: Prentice-Hall).

JÈHNE, B., 2005, *Digital Image Processing* (Berlin: Springer).

KRAAIKAMP, C., DEKKING, M., ELFEKI, A. and BRUINING, J., 2001, Multiscale and multi-resolution stochastic modeling of subsurface heterogeneity by tree-indexed Markov chains. *Computers Geoscience*, **5**, pp. 47–60.

KRESSLER, F.P., STEINNOCHER, K. and FRANZEN, M., 2005, Object-oriented classification of orthophotos to support update of spatial databases. In *IGARSS '05 Proceedings*, **1**, pp. 253–256.

LE MOIGNE, J. and TILTON, J., 1995, Refining image segmentation by integration of edge and region data. *IEEE Transactions on Geoscience and Remote Sensing*, **33**, pp. 605–615.

MARDIA, K. and HAINSORTH, T., 1988, A spatial thresholding method for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **10**, pp. 910–927.

PAVLIDIS, T. and LIOW, Y., 1990, Integrating region growing and edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **12**, pp. 225–233.

RUSS, J.C., 1999, *The Image Processing Handbook* (Boca Raton, FL: CRC Press), pp. 371.

SHACKELFORD, A.K. and DAVIS, C.H., 2003, A combined fuzzy pixel-based and object-based approach for classification of high-resolution multispectral data over urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, **41**, pp. 2354–2363.

SHANDLEY, J., FRANKLIN, J. and WHITE, T., 1996, Testing the Woodcock–Harward image segmentation algorithm in an area of southern California chaparral and woodland vegetation. *International Journal of Remote Sensing*, **17**, pp. 983–1004.

WANG, X. and WANG, H., 2004, Markov random field modeled range image segmentation. *Pattern Recognition Letters*, **25**, pp. 367–375.

WANG, Y., WEI, W., ZHAO, S. and CHEN, X., 2004, Object-oriented classification and application in land use classification using SPOT-5 PAN imagery. In *IGARSS '04. Proceedings*, **5**, pp. 3158–3160.

WEI, W., CHEN, X. and MA, A., 2005, Object-oriented information extraction and application in high-resolution remote sensing image. In *IGARSS '05. Proceedings*, **6**, pp. 3803–3806.

WORBOYS, M. and DUCKHAM, M., 2004, *GIS: A Computing Perspective* (Boca Raton, FL: CRC Press).