A SUPERVISED APPROACH FOR THE ESTIMATION OF PARAMETERS OF MULTIRESOLUTION SEGMENTATION AND ITS APPLICATION IN BUILDING FEATURE EXTRATION FROM VHR IMAGERY

VIVEK DEY

September 2011

GEODESY AND GEOMATICS ENGINEERING
UNB

TECHNICAL REPORT NO. 278

A SUPERVISED APPROACH FOR THE ESTIMATION OF PARAMETERS OF MULTIRESOLUTION SEGMENTATION AND ITS APPLICATION IN BUILDING FEATURE EXTRACTION FROM VHR IMAGERY

Vivek Dey

Department of Geodesy and Geomatics Engineering
University of New Brunswick
P.O. Box 4400
Fredericton, N.B.
Canada
E3B 5A3

September, 2011

PREFACE

This technical report is a reproduction of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Engineering in the Department of Geodesy and Geomatics Engineering, September 2011. The research was supervised by Dr. Yun Zhang and Dr. Ming Zhong, and support was provided by the Natural Sciences and Engineering Research Council of Canada.

 The citation to this work should appear as follows:

Dey, Vivek (2011). A *Supervised Approach for the Estimation of Parameters of Multiresolution Segmentation and its Application in Building Feature Extraction from VHR Imagery*. M.Sc.E. thesis, Department of Geodesy and Geomatics Engineering Technical Report No. 278, University of New Brunswick, Fredericton, New Brunswick, Canada, 162 pp.

DEDICATION

My thesis is dedicated to my Uncle Mr. M.L. Goel. You were the source of energy for this work.

ABSTRACT

With the advent of very high spatial resolution (VHR) satellite, spatial details within the image scene have increased considerably. This led to the development of object-based image analysis (OBIA) for the analysis of VHR satellite images. Image segmentation is the fundamental step for OBIA. However, a large number of techniques exist for RS image segmentation. To identify the best ones for VHR imagery, a comprehensive literature review on image segmentation is performed. Based on that review, it is found that multiresolution segmentation, as implemented in the commercial software eCognition™, is the most widely-used technique and has been successfully applied for a wide variety of VHR images. However, multiresolution segmentation suffers from the parameter estimation problem. Therefore, this study proposes a solution to the problem of the parameter estimation for improving its efficiency in VHR image segmentation.

The solution aims to identify the optimal parameters, which correspond to effective segmentation. The solution to the parameter estimation is drawn from the equations related to the merging of any two adjacent objects in multiresolution segmentation. The solution utilizes spectral, shape, size, and neighbourhood relationships for a supervised solution. In order to justify the results of the solution, a global segmentation accuracy evaluation technique is also proposed. The solution performs excellently with the VHR images of different sensors, scenes, and land cover classes.

In order to justify the applicability of solution to a real life problem, a building detection application based on multiresolution segmentation from the estimated parameters, is carried out. The accuracy of the building detection is found nearly to be eighty percent. Finally, it can be concluded that the proposed solution is fast, easy to implement, and effective for the intended applications.

ACKNOWLEDGEMENTS

I would primarily thank my supervisor Dr. Yun Zhang and Dr. Ming Zhong for providing the opportunity to conduct this research. Without the support, the research would not have been possible. In addition, I would like to thank Dr. Yun Zhang for supervising, resources, and proper guidance to complete the research and this thesis.

I would also like to thank the City of Fredericton, ISPRS, GeoEye®, and DigitalGlobe® for providing the data used in this research. The wide variety of data was essential to this study.

I would also to extend my thanks to the staff and colleague at the GGE Department, especially Sylvia Whitaker and David Fraser, for helping me out with the software and the academic procedures at UNB. The study environment and the supporting staff were excellent and I have made several life long bonds in the department.

I would also like to thank my parents, my sister, and my dearest friend, Anshu, for giving me a boost to complete the thesis in time.

Last, but not the least, I would like to thank my deceased Uncle, M. L. Goel, who always gave me hope for moving ahead during my research's tough times.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ANN – Artificial Neural Network

B – Blue

BF – Branching Factor

DP – Detection Percentage

FbSP – Fuzzy based Segmentation Parameter

FCM – Fuzzy C-Means

FIRME – Fuzzy Image Region Method

FIS – Fuzzy Inference System

FNEA – Fractal Net Evaluation Approach

FP – False Positive

G – Green

GML – Gaussian Maximum Likelihood

GRF – Gibbs Random Field

HSMR – Hierarchical Split and Merge

IDL – Interactive Data Language

LiDAR – Light Detection and Ranging

MAP – Maximizing a Posterior

MC – Marker Controlled

MDN – Mean-Difference to Neighbours

MOSA – Multi-scale Object Specific Analysis

MPM – Maximizing Posterior Marginal

MR – Multiresolution

MRF – Markov Random Field

MS – Multispectral

MSc – Multi-scale

NDVI – Normalized Difference Vegetation Index

NIR – Near Infra-red

OBIA – Object-based Image Analysis

OWO – Objects without Over-segmentation

OWU – Objects without Under-segmentation

Pan – Panchromatic

PGF – Peer-Group Filter

R – Red

RISA – Region based Image Segmentation

RS – Remote Sensing

SCRM – Size-Constrained Region Merging

SS – Scale Space

TN – True Negative

TP – True Positive

VHR – Very High Resolution

CHAPTER 1

INTRODUCTION

This research presents a methodology to enhance the segmentation results of very high spatial resolution (VHR) remote sensing (RS) satellite imagery. The research identified the multiresolution segmentation as the most widely-used RS image segmentation techniques in a comprehensive literature review. Further, the research proposed a methodology to estimate the optimal parameters of the multiresolution segmentation in order to improve the segmentation results. The methodology incorporates image segmentation theories, remote sensing principles, and mathematical analysis. This paper-based M.Sc.E. thesis is presented through the following papers:

<u>Paper1 (peer reviewed):</u>

Dey, V., Y. Zhang, M. Zhong, and B. Salehi (2011). "Image Segmentation Techniques for Urban Land Cover Segmentation of VHR Images: Recent Developments and Future Prospects." *International Journal of Remote Sensing,* (under review).

<u>Paper2 (peer reviewed):</u>

Dey, V., Y. Zhang, and M. Zhong (2011). "A Supervised Methodology for Optimal Parameter Estimation of Multiresolution Segmentation within eCognition." *International Journal of Remote Sensing,* (to be submitted).

<u>Paper3 (peer reviewed):</u>

Dey, V., Y. Zhang, and M. Zhong, and B. Salehi (2011). "Building Detection using Multi-Level Segmentation with a Fuzzy Parameter based Region Merging Criteria." Proceedings of the *32nd Canadian Symposium on Remote Sensing,* Sherbrooke, Québec, Canada, 13-16 July 2011, pp. 1-8.

1.1 Thesis Structure

This thesis follows the structure of a paper-based thesis. The thesis incorporates two journal papers (to be submitted for peer-review) and one conference paper (published). The thesis includes five chapters: introduction, three research papers (each as one chapter), and conclusions. The chapters are followed by Appendix I, which states the mathematical derivations used for the analysis of this work. Figure 1.1 illustrates the organization of this thesis.

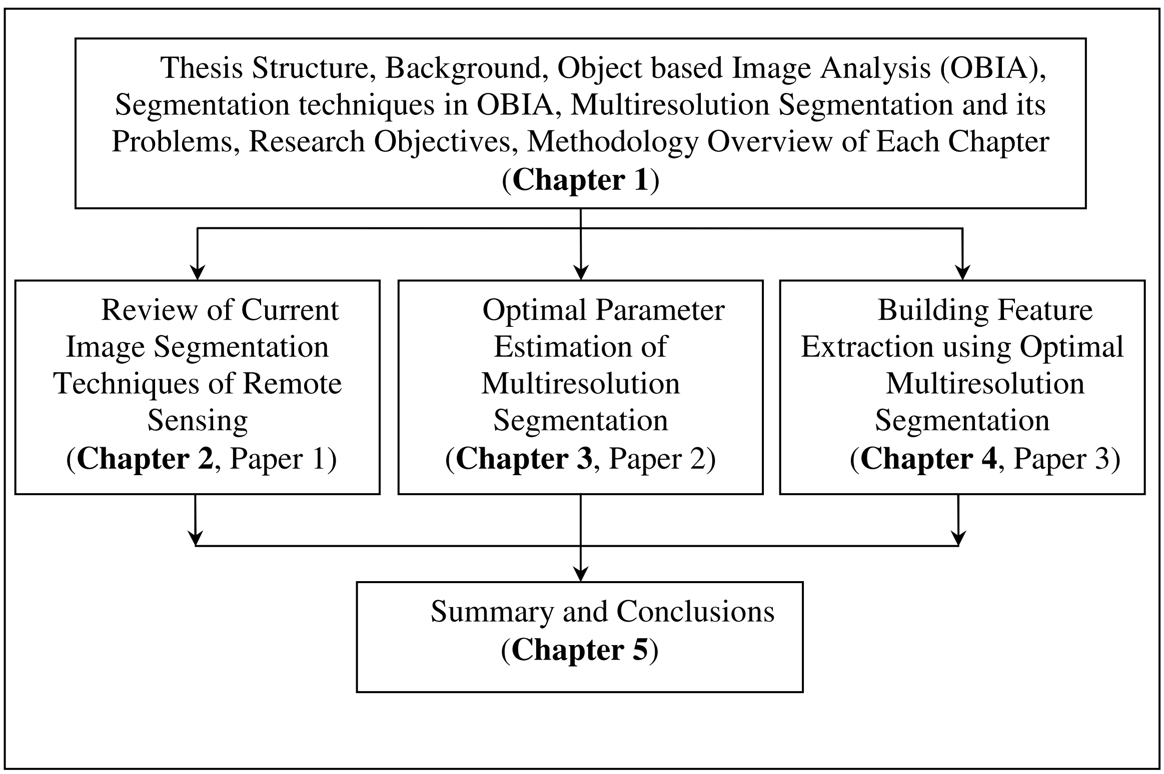


Figure 1.1: Organization of this thesis

1.2 Background

1.2.1 The Era of VHR Satellite Imagery

The era of VHR satellite imagery (spatial resolution <= 1m) began with the successful launch of IKONOS-2 in 1999. In the last decade, the spatial resolution of satellite images has continuously improved. For example, the maximum spatial resolution of panchromatic (Pan) band of GeoEye-1 satellite is 0.41m and 2.44m for each of its four multi-spectral (MS) bands: red (R), Green (G), blue (B), and near infra-red (NIR). The requirement of an appropriate signal to noise ratio for the high spatial resolution data collection has restricted the spectral resolution to four bands for a decade since 1999 [Zhang, 2004]. However, the successful launch of WorldView-2 with 8 MS bands (each with a spatial resolution of 1.8m) and 1 Pan band (with maximum spatial resolution of 0.46m) has relaxed this restriction. Table 1.1 summarizes the development of optical VHR satellite images since 1999. The development of VHR satellite imagery is attributed to the two factors: (1) the growth in sensor technology; and (2) the change in US government policy towards commercially available spatial resolution of the images from remote sensing satellites.

Table 1.1: The growth of optical VHR satellite images since 1999.

Sensor Type	Spatial Resolution (in metres)		Number of Bands	Year	Country
	MS	Pan			
IKONOS-2	4.0	1.0	4MS + 1Pan	1999	US
Quickbird-2	2.44	0.61	4MS + 1Pan	2001	US
Kompsat-2	4	1.0	4MS + 1Pan	2006	South Korea
EROS-B	0.7		1Pan	2006	Israel
CartoSat-2	-	0.8	1Pan	2007	India
WorldView-1	-	0.5	1Pan	2007	US
Skymed-1		1.0	1Pan	2007	Italy
Skymed-2		1.0	1Pan	2007	Italy
GeoEye-1	1.65	0.41	4MS + 1Pan	2008	US
Skymed-3		1.0	1Pan	2008	Italy
WorldView-2	1.8	0.46	8MS + 1Pan	2009	US
Skymed-3		1.0	1Pan	2010	Italy
Pléiades-1	2.0	0.5	4MS + 1Pan	2011	Italy
Pléiades-2	2.0	0.5	4MS + 1Pan	2012	France
Kompsat-3	3.2	0.7	4MS + 1Pan	2012	South Korea
GeoEye-2	1.0	0.25	4MS + 1Pan	2013	US

Pan = Panchromatic, MS = Multi-Spectral, Source: [Stoney, 2008].

Undoubtedly, VHR satellite imagery is providing high amount of data related to earth observation every day. This is because satellite images have wide ground coverage as well as high frequency of image acquisition [Konecny and Schiewe, 1996]. Figure 1.2 depicts the improvement in the visual details of ground objects of the selected area with the increased spatial resolution. This improvement enabled the use of VHR satellite imagery in a rapidly growing list of new applications, e.g., urban security, urban disaster management, and urban planning [Blaschke, 2010].

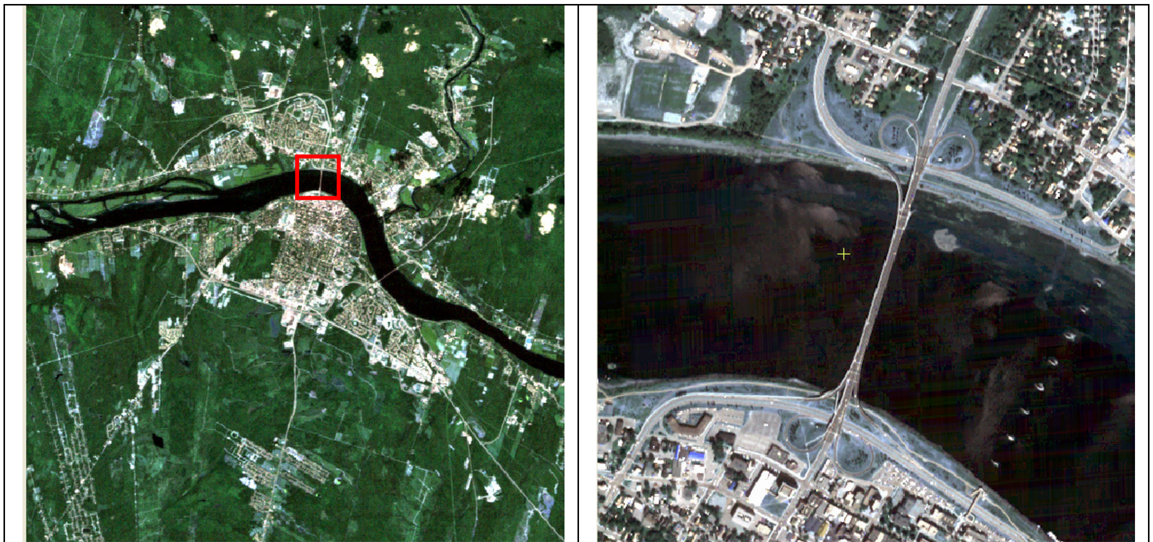

[a] [b]

Figure 1.2: Depiction of the visual details in low and high resolution satellite images of same dimension, 512x512 pixels: [a] Landsat TM 30m MS and [b] Quickbird 2.44m MS. Source: Wuest [2009].

1.2.2 Development of Object-based Image Analysis for VHR image

Although a VHR image can be best analyzed by a human interpreter, the growing databases of VHR satellite images have necessitated a computer-based VHR image analysis to assist in the human-based interpretation. A computer-based analysis of a VHR image should utilize major features of a VHR image to match the level of human analysis. The major features of a VHR image are:

1. The high spectral variance within the land covers classes of a VHR image. For example, on visual analysis of Figure 1.2[a], the urban regions in the selected area (in red) of the low resolution image look spectrally more

homogeneous as compared to the same area of the high resolution satellite image shown in Figure 1.2[b]

2. The ability to identify the small objects (e.g., small residential buildings, swimming pools, parking lots, and trees) along with the large objects (e.g., commercial buildings, lakes, roads, and forests) from the same image.

The traditional pixel-based analysis failed to utilize these features because the pixel size of VHR imagery is too small to depict the structure of common land cover classes such as forests, urban areas, and water-bodies [Woodcock and Strahler, 1987; Castilla and Hay, 2008]. Further, pixel-based analysis was unsuccessful in achieving better results using VHR images for common applications, e.g., land cover classification, feature extraction, and change detection [Blaschke and Strobl, 20001; Flanders et al., 2003; Benz et al., 2004; Blaschke et al., 2006; Zhou et al., 2008; Linli et al., 2008]. This failure led to the surge of object-based image analysis (OBIA) for VHR imagery [Hay et al., 2003; Hay and Castilla, 2006; Blaschke, 2010].

The term "objects" in OBIA refers to the groups of pixels which form contiguous regions in an image [Benz et al., 2004]. In addition, the contiguous regions should have a sense of uniformity (e.g., spectral uniformity and textural uniformity) as well as contrast with the surroundings such that the region can be identified as a distinct entity [Castilla and Hay, 2008]. Further, in RS, the object should correspond with a geographical entity (e.g., a tree, a building, and a lake). Figure 1.3 illustrates the concept of meaning of object as a geographical entity using GeoEye-1 image of Hobart, Tasmania, Australia.

Overall, the aim of OBIA is to identify the meaningful objects and link them to associated geographical classes [Hay and Castilla, 2006; Blaschke, 2010]. In simple words, OBIA tries to imitate the human interpretation of RS images as efficiently as possible through computer-based procedures [Castilla and Hay, 2008].

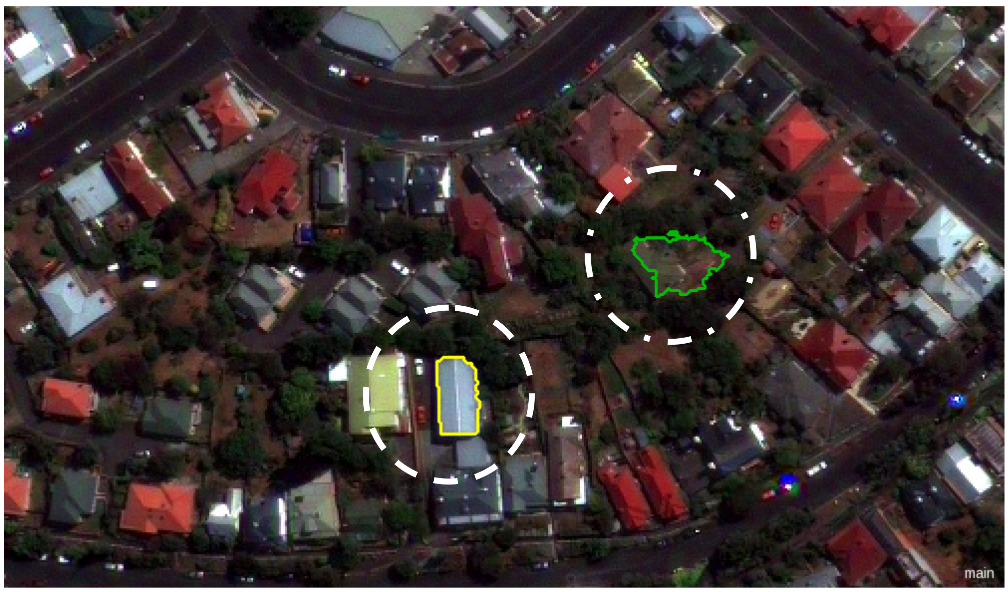


Figure 1.3: Illustrates the image object (encircled in white with yellow outline) which corresponds with a geographical entity or class (building) whereas the image object (encircled in white with green polygon outline) does not correspond with a geographical entity. (Source: GeoEye-1 image of Hobart, Australia, ©GeoEye and ISPRS.

1.2.3 Segmentation Techniques Required for OBIA

The first step of OBIA is to identify the image objects using computer-based digital image processing. As per the definition of the image objects, *image segments*, as the outputs of a computer-based image segmentation procedure, closely resemble the image objects. In the image segmentation process, the image is partitioned into homogeneous segments (also known as regions) identified based on a homogeneity criterion, such that each segment is distinct from its surrounding segments as per the homogeneity criteria [Pal and Pal, 1993; Blaschke, 2010]. Therefore, the first step of OBIA involves to identifying image objects using an appropriate image segmentation technique.

Although a wide variety of image segmentation techniques exist for single band grey-level images in Computer Vision and Bio-medical image analysis, only few are appropriate for RS purposes due to the requirement of multi-scale and multi-band analysis in RS [Cheng et al., 2001; Castilla and Hay, 2008]. This is further justified by the fact that the appropriate image segmentation techniques differs depending upon the desired outcome of an OBIA applications [Pal and Pal, 1993; Blaschke et al., 2006]. However, since 2001, a substantial research has been devoted to RS image segmentation techniques for OBIA [Shankar, 2007; Blaschke, 2010]. The major categories of segmentation techniques identified in this thesis are: (1) Clustering, (2) Level-set models, (3) Markov-Random fields, (4) Fuzzy logic based techniques, (5) Neural network based techniques, (6) Multi-scale techniques, (7) Watershed model, and (8) Hierarchical split and merge model [Shankar, 2007; Dey et al., 2010]. A detailed description of each category and their techniques applied for the VHR image segmentation is provided in Chapter 2.

In VHR RS image analysis, it has been widely accepted that the multi-scale based techniques are the best among all the above-mentioned categories of segmentation techniques [Hay et al., 2003; Blaschke et al., 2006; Blaschke, 2010]. Moreover, it has been found that multiresolution segmentation, a multi-scale segmentation technique, has been successfully and widely employed in most of the VHR image applications, e.g., urban buildings detection, forest management, disaster management, animal habitat identification, and glacier mapping [Blachke, 2010]. Therefore, this thesis has selected multiresolution segmentation in its processing of VHR image application to support residential building detection (described in Chapter 4).

1.3 Multiresolution Segmentation and its Problems

Multiresolution segmentation, proposed by Baatz and Schäpe [2000], is a multi-scale region-based segmentation technique where the different scales are used for the analysis of objects of different sizes [Benz et al., 2004]. It uses spectral, shape, and size features to identify the criteria of region merging based segmentation. Being a region-based technique, multiresolution segmentation is a bottom-up approach, where a seed pixel is selected to grow a region and the seeds are distributed uniformly over the image for a parallel growth of regions [Benz et al., 2004]. The different scales required for multi-scale segmentation are represented with the hierarchically connected segmentation levels, where the segments of different sizes are generated according to the specified scale parameters of each level [Benz et al., 2004]. Figure 1.4 illustrates the two different segmentation levels and their scales representing the different sizes of segments. From

Figure 1.4, it is clear that the scale parameter has high impact on the meaningful object detection from the VHR imagery. Apart from the scale parameter, multiresolution segmentation requires two more parameters: shape weight and the compactness weight [Definiens AG, 2009]. Hence, a successful multiresolution segmentation requires estimation of the three optimal parameters.

[a] Level 1 segmentation with the scale-75, shape-0.84, compactness-0.77

[b] Level 2 segmentation with the scale-327, shape-0.76, compactness-0.74.

Figure 1.4: Illustrates the two hierarchical level of segmentation, where at [a] meaningful segmentation of urban forest (in red) has not been achieved and at [b] it is achieved.

1.3.1 The Problems

This Section describes some of the common problems faced in the general segmentation of VHR imagery and how the multiresolution segmentation tackles them. The general problems of the segmentation are:

1. Estimation of the optimal parameters.

2. Reproduction of the segmentation results given the same segmentation conditions.
3. Inclination on operator.
4. Under-segmentation and over-segmentation.
5. Execution time of the algorithm

1.3.1.1 Optimal Parameter Estimation

Similar to most of the VHR image segmentation techniques, multiresolution segmentation requires three operator-dependent parameters. These parameters control the segmentation results to a high extent [Möller et al., 2007; Tian and Chen, 2007]. The traditional procedure of identifying parameters is by trial and error procedure, where the parameters are arbitrarily assigned to segment the image until an operator-based visually pleasing segmentation results are achieved. Fortunately, the multiresolution segmentation has been implemented in a commercial software eCognition™ (now owned by Trimble Inc.), which has made the trial and error approach relatively easy to follow [Benz et al., 2004; Definiens AG, 2009]. However, the traditional procedure is still very tiresome and may result in a sub-optimal solution due to the subjective nature of the visual evaluation of the segmentation results [Hay et al., 2003; Marpu et al., 2010]. Hence, optimal parameter estimation is one of the major challenges to achieve an efficient multiresolution segmentation.

1.3.1.2 Reproduction of the Segmentation Results

Often, the segmentation techniques fail to reproduce the same results even with the same parameter settings [Baat and Schäpe, 2000]. This is primarily due to the process of random seed distribution required for the parallel processing of region growing across the image [Baatz and Schäpe, 2000; Benz et al., 2004]. Moreover, most of the segmentation techniques produce different segmentation results for the same region from the different subsets [Tian and Chen, 2007]. Multiresolution segmentation excels in solving both the above-mentioned problems with its proprietary algorithm related to the parallel distribution of seed-pixels for region-growing. Hence, the segmentation results of the multiresolution segmentation are reproducible: (1) for the same area from different subsets, irrespective of the size of the subset; and (2) for the same set of parameters.

1.3.1.3 Inclination on Operator

In image segmentation, the estimation of parameters and the assessment of the segmentation results depend on the operator. While optimal parameter estimation depends upon the operator's experience with the segmentation technique, the accuracy assessment requires familiarity with the geographical area of the imagery. The parameter estimation problem has already been discussed in the Section 1.3.1.1. In order to minimize the operator's dependency, several segmentation assessment techniques have been proposed. However, in spite of the several proposed quantitative assessment techniques for the optimal segmentation results, *visual* assessments are still widely used [Lang et al., 2009; Corcoran et al., 2010; Dey et al., 2010]. The assessment of the results of multiresolution segmentation is also subjective based on each individual operator. Hence, a comprehensive objective assessment, which can remove or minimize the operator dependency as well as easy to implement and analyze, is required for efficient multiresolution segmentation.

1.3.1.4 Under-segmentation and over-segmentation

Under-segmentation and over-segmentation are specific to the result assessment and help in defining the optimal segmentation results (required for optimal parameter estimation). Kim et al. [2008] defined the scale of optimal segmentation as "the one that is not over-segmented, with an excessive number of segments that are on average too small, and also not under-segmented, with too few segments that are on average too large." Similarly, Castilla and Hay [2008] stated that "a good segmentation is one that shows little over-segmentation but no under-segmentation". Often, an over-segmentation is preferred over under-segmentation because unlike under-segmentation, the over-segmentation problem can be handled in the subsequent analysis of segmentation (e.g., segment-based classification) [Castilla and Hay, 2008]. The identification of over-segmentation and under-segmentation can be based on the reference segment as well as the operator's visual assessment, as illustrated in Figure 1.5. Hence, the objective accuracy assessment should be based on an optimal segmentation, which minimizes both the under-segmentation as well as the over-segmentation.

Figure 1.5: Illustrates the over-segmentation [a], correct segmentation [b], and under-segmentation of a building object (in red), identified at [b].

1.3.1.5 Execution Time of Algorithm

With the rapid increase in the computational power of the computers, the execution time has become a less significant factor. The time complexity is mostly related to the design of the computer-based algorithms. Multiresolution segmentation has been efficiently implemented in commercial software, eCognition™ since 1999 (now eCognition™ Developer 8). This implementation has facilitated the use of the software among researchers and in industry. However, if a large size image is segmented with a very low scale parameter value, then the segmentation results might have some memory problems [Definiens AG, 2009]. Overall, the problem of time complexity has been solved for the case of multiresolution segmentation through commercial implementation.

From the discussions in the last five Sections, the major problem is found to be the parameter estimation. In order to solve the problem of parameter estimation, two more problems have been identified. Hence, the three major problems of the multiresolution segmentation are related to:

1. The estimation of parameters;
2. The identification of criteria leading to optimal segmentation; and
3. The identification of objective segmentation accuracy assessment technique.

1.4 The Objective

The objective of this research is to present a methodology to enhance the segmentation results of VHR imagery. In order to achieve the objective, the specific tasks are:

1. To identify the widely-used segmentation techniques and select the best among them.
2. To determine the problems in the selected widely-used technique, i.e., multiresolution segmentation.
3. To propose a solution to the problems of the multiresolution segmentation.

4. To justify the proposed solution of the problems by the means of the experiments on different VHR images.
5. To apply the multiresolution segmentation, with the estimated parameters, for the building feature extraction application.

1.5 The Methodology

This Section gives a brief description of the research performed for accomplishing the objectives of this study. The research is conducted on the images of different locations and sensors. Table 1.2 enlists the description of the images used for this research, where each MS image has: Red (R), Green (G), Blue (B), and Near Infrared (NIR) spectral bands. In order to achieve the high spatial resolution for analysis, each of the MS images is pansharpened using its corresponding Pan image to retain the spectral resolution of MS images as well as the spatial detail of the Pan image. Figure 1.6 illustrates the methodology followed in this research. The components of the methodology are as follows:

1. To complete the first task, an extensive literature review is carried out on various image segmentation techniques used for the urban land cover segmentation of VHR image. After the literature review, the multi-scale and the watershed based segmentation techniques are identified as the

most widely-used techniques and the multiresolution segmentation, a multiscale technique, is selected for this research.

2. To complete the second task related to the specified problems, a literature review of applications of multiresolution segmentation is carried out to identify the problems (optimal parameter estimation and its sub-problems: definition of optimal segmentation and accuracy assessment) of multiresolution segmentation.
3. To achieve the third task, the performed steps are:
    a. Review of the existing solutions to the parameter estimation.
    b. Identification of the features required for the estimation of the optimal parameters.
    c. Implementation of the proposed approach using functionalities of eCognition™ for faster processing.
4. To finish the fourth task, the three VHR images (specified in Table 1.2) are employed for the segmentation experiments using the estimated parameter.
5. To achieve the final task of the building extraction, several characteristic features of the buildings (e.g., building edge, building shape, building shadow, and rooftop color homogeneity) are identified. These characteristic features are utilized to aid in the detection of building segments, which are resulted from multiresolution segmentation.

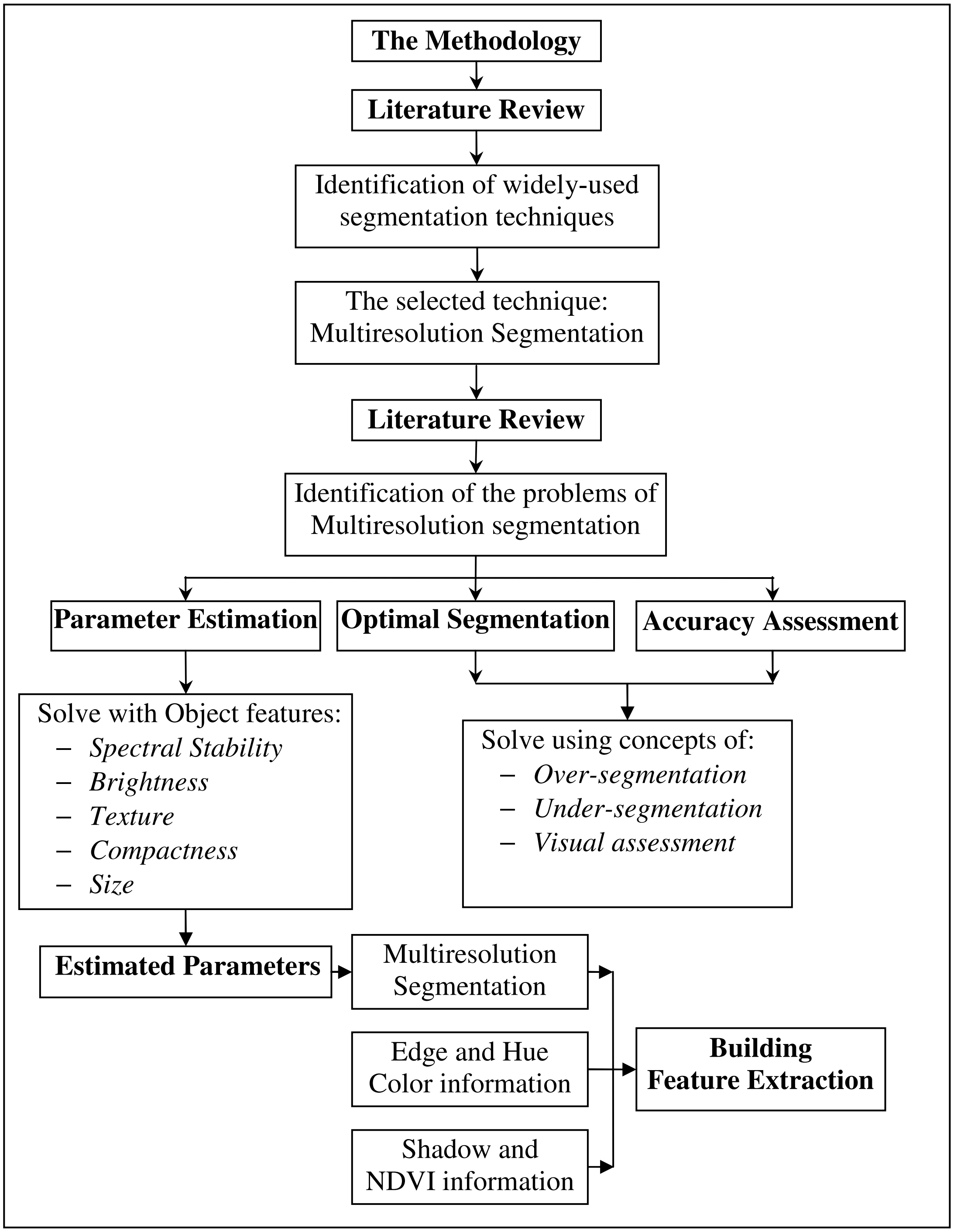


Figure 1.6: Illustrates the methodology of the research of this thesis. The bold text states the major steps of the methodology and the rest as the outcomes or inputs of the steps.

Table 1.2 The images used for the experiments of this research in different chapters

Image	Bands used	Resolution	Area	Year	Used Section
IKONOS	4 MS + 1Pan	Pan 1.0m MS 4.0m	Fredericton, New Brunswick, Canada	2002	Chapter 3
Quickbird	4 MS + 1Pan	Pan 0.61m MS 2.44 m	Fredericton, New Brunswick, Canada	2002	Chapter 3
GeoEye-1	4 MS + 1Pan	Pan 0.5m MS 2.0m	Hobart, Tasmania, Australia	2009	Chapter 3 & 4

1.5.1 Review of the Existing Solutions

This Section gives a review of the solutions to the problems of the multiresolution segmentations. The three problems (specified in the Section 1.3.1) are inter-related; hence, the solutions are also inter-related. Maxwell [2005] was the first to propose the solution using a supervised fuzzy approach. He used spectral, shape, size, texture, and neighbourhood features to identify the optimal parameters but the assessment technique was visual. Costa et al. [2008] proposed a genetic algorithm based solution, where few reference segments, are used to optimize the parameters. Tian and Chen [2007] and Marpu et al. [2010] specified their objective evaluation measures for identifying the optimal segmentation from a sequence of parameter settings. Drăgut et al. [2010] used auto-correlation among the segments to identify the optimal scale parameter. This thesis has utilized the object information/features similar to the object information of Maxwell [2005] (see Figure 1.7) for the solution of the problems of multiresolution segmentation.

1.5.2 The Proposed Solution for the Multiresolution Segmentation

The proposed solution is a heuristically derived supervised approach. The building blocks of the proposed solution are based on the:

1. Features of an object and the supervised approach of Maxwell [2005], also known as Fuzzy based supervised approach (FbSP) [Zhang et al., 2010]. The features are: (a) spectral standard deviation, (b) spectral mean, (c) mean-difference to neighbours (Spectral stability), (d) compactness, and (f) size of the segments.
2. Optimal segmentation conditions stated by Castilla and Hay [2008] and Kim et al., [2008] (see Section 1.3.1.4).
3. The objective accuracy assessment techniques specified by Möller et al. [2007], Tian and Chen [2007], and Marpu et al. [2010].

1.5.2.1 The FbSP Solution

FbSP approach was proposed by Maxwell [2005]. FbSP is a fuzzy logic approach, which is applied in a supervised manner. Each of the three parameters (Scale, Shape weight, and Compactness weight) is estimated using different fuzzy inference networks (FIS) namely, Scale FIS, Shape FIS, and Smoothness FIS (Smoothness weight = 1-Compactness weight) [Tizhoosh and Haussecker, 2006]. FbSP uses features of a training object of a particular class (e.g., forest object in Level 2 of Figure 1.4[b]) and its over-segmented sub-objects (forest over-segmentation of Level 1 of Figure 1.4[a]) to obtain the three parameters. Figure 1.6 illustrates the workflow of the FbSP.

1.5.2.2 The Objective Assessment of Segmentation Accuracy

The segmentation results are assessed using the conditions of over-segmentation and under-segmentation by Tian and Chen [2007] as well as Marpu et al. [2010]. However, the comparison for the assessment relies on few reference segments. Such a comparison is essentially a local assessment technique. In this thesis, a global assessment process has been proposed. The proposed assessment process uses the overlap percentage between the reference segments and the obtained segments to derive the amount of over-segmentation and under-segmentation.

Based on the proposed assessment technique, the segmentation results obtained from multiresolution segmentation have been found to be quantitatively comparable to the results of the tradition trial and error based solution. This justified that the results obtained from the estimated parameters are effective. However, the tradition approach may require hours for the estimation whereas the proposed solution provides the effective parameters always within 5 minutes. Therefore, the major **contribution of the** proposed solution is this reduction of time in estimation of the effective parameters of multiresolution segmentation.

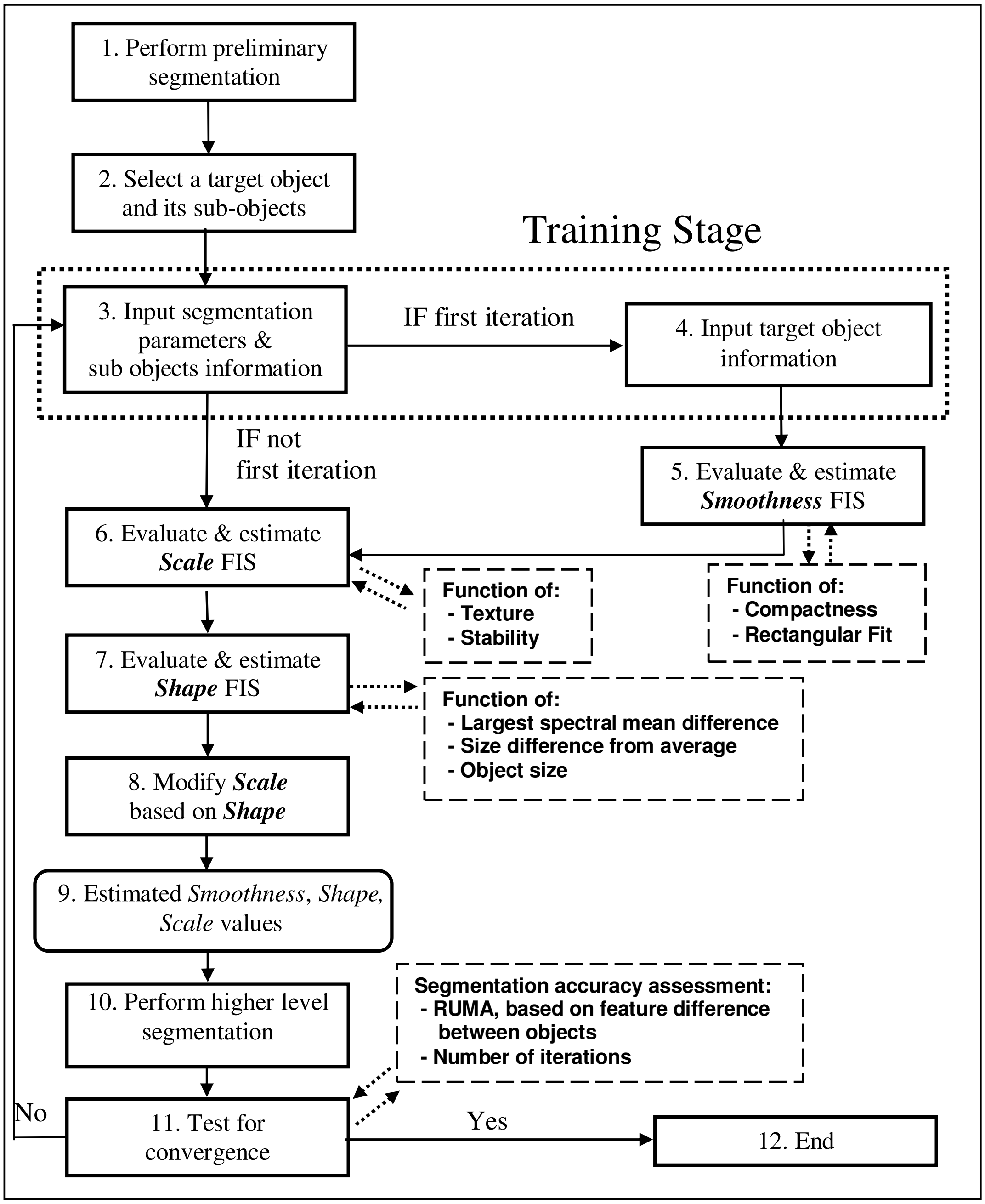


Figure 1.7: Workflow of the proposed FbSP optimizer. The values of the current Segmentation Parameters (Smoothness (1-$w_{compact}$), Shape (1-w) and Scale (s)), Sub objects information (Texture, Stability, Brightness, and Area) and Target Object information (Texture, Stability, Brightness, Area, Rectangular Fit, and Compactness) are inputted into FbSP optimizer to train the FISs (Fuzzy Inference Systems) to estimate the optimal Segmentation Parameters for the Target Object in an iterative process. Source: Maxwell [2005].

1.5.3 Building Extraction

This Section gives the outline of the methodology followed to achieve the final objective: building extraction from the VHR imagery using multiresolution segmentation. Building extraction is one of the basic operations of any urban applications (e.g., urban mapping, urban change detection, urban sprawl, and urban security) and has been studied for more than two decades [Lin and Nevatia, 1998; Luan and Ye, 2010]. Building extraction is primarily performed using aerial images [Lin and Nevatia 1998; Mayer, 1999; Luan and Ye, 2010]. It is because aerial images are acquired as stereo-pairs, capable of providing building height data, which is crucial information for building detection. Apart from the aerial imagery, LiDAR can also be used for the height information [Rottensteiner et al., 2003]. However, with the development of VHR images, the VHR satellite images are being increasingly used for the building extraction [Luan and Ye, 2010]. According to Konecny and Schiewe [1996] and Baltsavias [2004], this is primarily due to the following:

1. Increased spatial details of VHR image, which led to identification of the complexity of the land cover classes. For example, different trees' canopies in a forest and small residential buildings in urban area (see Figure 1.2).
2. Extended ground coverage of a VHR satellite image compared to an aerial image. Hence, a VHR image acquisition is much cheaper than the acquisition of an aerial image or LiDAR data. For example, a single scene

of GeoEye-1 covers around 15 Km x 15 Km whereas a single scene of HRSC-A covers around 0.48 Km x 0.48 Km.

3. Acquired at the time (around 10 to 11:00 AM) when the sun's rays cast shadows of the building, a major cue for the identification of the buildings.

1.5.3.1 Review of the Existing Building Extraction

Apart from the utilizing 3D height feature, the basic principles used for the building extraction using VHR imagery are similar to the aerial imagery. Mayer [1999], Baltsavias [2004], and Luan and Ye [2010] reviewed several building extraction approaches using only 2D information. This thesis employs the categorization of Mayer [1999] for reviewing the building extraction approaches. The categorization comprises: (1) model-based approaches; and (2) strategy-based approaches, described next.

1.5.3.1.1 Model-based approaches

Model based methods utilize modelling of the following information:

(1) Shape of the buildings, e.g., rectangular, L-shaped, defined prior shape (see Figure 1.3 for different shapes) [Lin and Nevatia, 1998; Song et al., 2006; Sirmaçek and Ünsalan, 2008; Karantzalos and Paragios, 2009].

(2) Statistics based modelling, e.g., active contour modelling or level sets [Peng et al., 2005; Mayunga et al., 2007; Karantzalos and Paragios, 2009].

Moreover, a hybrid of two or models are also possible. For example, Karantzalos and Paragios [2009] used shape as well as level set based models for building detection using a VHR urban image. Often models based approaches fail because of their lack of generalization across VHR images of different sensors and locations [Luan and Ye, 2010]. The failure is attributed to following reasons [Mayer, 1999]:

1. Shape models will fail if the VHR image has complex building-rooftops. For example, Lin and Nevatia [1998] assumed rectangular shape of the building, which may not be true for the complex buildings shapes, such as those found in Figure 1.3.
2. Statistical based models are complex and lack incorporation of human knowledge, which is beneficial for the building extraction [Baltsavias, 2004].

However, model-based approaches provide a direct vector output for the extracted buildings, which is a huge advantage. Hence, the approaches are widely-popular for image specific customized solutions.

1.5.3.1.2 Strategy-based approaches

Strategy-based building detection approaches capture the general properties (e.g., contrast, texture, and context) of an image to aid in building extraction. The most widely used strategy-based approach is segmentation-guided feature extraction. Shackelford and Davis [2003] performed segmentation and used shadow context to identify buildings from the segmented image. Muller and Zaum [2005] and Song et al. [2006] also used region-based segmentation, shadow context, and topological relationships for building extraction. Strategy-based approaches are superior in the sense that they can utilize human knowledge in their interpretation and can be generalized for VHR images covering different locations. However, due to the lack of customization of the strategy-based approaches, they provide lower building extraction accuracies compared to the model-based approaches.

Most of the model-based and strategy-based approaches utilize one or more of the following features of the buildings: (1) shadow context; (2) edge information; (3) rooftop homogeneity; (4) shape; (5) human knowledge (e.g., road near a building); and (6) texture. Therefore, the utilization of these features should improve the results of building extraction from a VHR image.

1.5.3.2 The Proposed Building Extraction Approach

The proposed approach is a strategy-based approach, where the general properties/features of buildings (specified in the Section 1.5.3.1.2) are utilized. The steps of the proposed approach are:

1. Perform an efficient multiresolution segmentation using the estimated parameters along with Sobel edge image and Hue image from Hue Saturation and Intensity Transform of RGB spectral bands of the VHR image [Richards and Jia, 2006]
2. Utilize shadow context for extracting the objects with the height information, i.e., trees and buildings.
3. Utilize Normalized Difference Vegetation Index (NDVI) features to separate the trees from the buildings.
4. Perform a knowledge-based Gaussian Maximum Likelihood (GML) classification to minimize the false detection of buildings.
5. Assess accuracy of the building extraction.

1.6 Overview of Each Chapter

The research of this thesis is presented in five chapters, where **Chapter 2, 3, and 4** comprise of the three research papers. Each chapter include Tables and Figures relevant to the topic being discussed.

As stated above, **Chapter 1** gives background information on: (1) development of VHR satellite imagery and OBIA; (2) the need of segmentation in OBIA; and (3) the problems of multiresolution image segmentation. In addition, it outlines the objectives of this research and the methodology followed to achieve the objectives. Finally, it outlines the content of each chapter to provide readers with the synopsis of each chapter.

The paper in **Chapter 2** reviews various digital image segmentation techniques available for VHR land cover segmentation of urban areas. It categorizes various segmentation techniques and identifies the widely-used techniques for VHR imagery. It concludes that the multi-scale segmentation techniques and watershed segmentation techniques are the most widely used techniques for urban land cover segmentation using VHR imagery.

The paper in **Chapter 3** proposes a solution to the problems related to the parameter estimation and the accuracy assessment of multiresolution segmentation. Multiresolution segmentation is identified as the most widely-used multi-scale segmentation technique based on the literature review of **Chapter 2**. In addition, it justifies the applicability of the proposed solution by the segmentation experiments on the VHR images of different resolutions, locations, and sensors.

The paper in **Chapter 4** applies the multiresolution segmentation for the building extraction from VHR imagery of Hobart, Australia. The aim of the building extraction is to access the performance of the segmentation results with the parameters, generated using the chosen parameter estimation approach. In addition, the Chapter assesses the accuracy of building feature extraction and recommends the ideas of further research.

The final Chapter, **Chapter 5**, summarizes the research presented in this thesis and draws conclusions based on the contribution of each chapter towards the goal of improvement of VHR image segmentation results. In addition, the areas of further research are suggested. The mathematical analyses for the solution proposed in **Chapter 3** are included in the **Appendix I**.

REFERENCES

Baatz, M. and A. Schape (2000). "Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation." *Angewandte Geographische Informations-Verarbeitung XII*. Ed. J. Strobl et al. AGIT Symposium, Salzburg, Germany, 2000, pp. 12-23. [Internet], cited 12 Aug. 2011. Available in .pdf format through www.ecognition.cc/download/baatz_schaepe.pdf

Baltsavias, E. P. (2004). "Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems."

ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 58, No. 3-4, pp. 129–151.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I. and M. Heynen (2004). "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information." *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 58, No.3-4, pp. 239-258.

Blaschke, T. and J. Strobl, (2001). "What's wrong with pixels? Some recent developments interfacing remote sensing and GIS." *GIS – Zeitschrift für Geoinformationssysteme,* Vol. 6, No. 1, pp. 12-17.

Blaschke, T., (2010). "Object based image analysis for remote sensing." *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 65, No. 1, pp. 2-16.

Blaschke, T., C. Burnett, and A. Pekkarinen (2006). "Image segmentation methods for object based analysis and classification." In *Remote Sensing Image Analysis: Including the Spatial Domain,* Ed. Jong, S.M. de and Meer F. D. van der. Springer-Verlag, Dordrecht, The Netherlands, pp. 211-236.

Castilla, G., and G. J. Hay (2008). "Image objects and geographic objects." In *Object-Based Image Analysis,* Ed. T. Blaschke, S. Lang and G.J. Hay. Springer-Verlag, Berlin: Hiedelberg, Germany, pp. 679-695.

Cheng, H. D., X. H. Jiang, Y. Sun, and J. Wang, (2001). "Color image segmentation: advances and prospects." *Pattern Recognition*, Vol. 34, No.12, pp. 2259-2281.

Corcoran, P., A. Winstanley, and P. Mooney (2010). "Segmentation performance evaluation for object-based remotely sensed image analysis." *International Journal of Remote Sensing,* Vol. 31, No. 3, pp. 617-645.

Costa, G.A.O.P., R. Q. Feitosa, T. B. Cazes, and B. Feijó (2008). "Genetic adaptation of segmentation parameters. "In *Object-Based Image Analysis,* Ed. T. Blaschke, S. Lang and G.J. Hay. Springer-Verlag, Berlin: Hiedelberg, Germany, pp. 679-695.

Definiens AG, 2009. *eCognition Developer 8.0, Reference Book*, published by Definiens AG. München, Germany.

Dey, V., Y. Zhang, and M. Zhong (2010). "A Review of image segmentation techniques with remote sensing perspective". Wagner W., and B. Székely (Eds.). *Proceedings of ISPRS TC VII Symposium- 100 Years ISPRS*, 5-7 July, Vienna, Austria, pp. 31-42.

Flanders, D., M. Hall-Beyer, and J. Pereverzoff (2003). "Preliminary Evaluation of eCognition Object-Based software for Cut Block Delinieation and Feature Extraction." *Canadian Journal of Remote Sensing*, Vol. 29, No. 4, pp. 441-452.

Hay, G. J. and G. Castilla (2006). "Object-Based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT)." In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science XXXVI* – 4/C42. 1st International Conference on Object-based Image Analysis (OBIA 2006), 4-5 July, Salzburg, Austria, pp. 1-3.

Hay, G.J., T. Blaschke, , D.J. Marceau, and A. Bouchard (2003). "A comparison of three image-object methods for the multiscale analysis of landscape structure." *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 57, No. 5-6, pp. 327-345.

Karantzalos, K., and N. Paragios (2009). "Recognition-driven two-dimensional competing priors toward automatic and accurate building detection." *IEEE Transactions on Geoscience and Remote Sensing*. Vol. 47, No. 1, pp. 133-144.

Kim, M., M. Madden, and T. Warner (2008). "Estimation of optimal image object size for the segmentation of forest stands with multispectral IKONOS imagery." In *Object-Based Image Analysis,* Ed. T. Blaschke, S. Lang and G.J. Hay. Springer-Verlag, Berlin: Hiedelberg, Germany, pp. 291–307.

Konecny, G., and J. Schiewe (1996). "Mapping from digital satellite image data with special reference to MOMS-02." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 51, No. 4, pp. 173–181.

Lang, S., E. Schöpfer, and T. Langanke (2009). "Combined object-based classification and manual interpretation-synergies for a quantitative assessment of parcels and biotopes." *Geocarto International*, Vol. 24, No. 2, pp. 99-114.

Lin C., and R. Nevatia (1998). "Building detection and description from a single intensity image." *Computer Vision and Image Understanding*, Vol. 72, No.2, pp. 101–121.

Linli C., S. Jun, T. Ping, and D. Huaqiang (2008). "Comparison study on the pixel-based and object-oriented methods of land-use/cover classification with TM data." In *IEEE: International Workshop on Earth Observations and Remote Sensing Applications*. EORSA 2008, 30 Jun-2 July, Beijing, China, pp. 1-5.

Luan, Q., C. Ye, and J. Long (2010). "Review of building information extraction based on single high-resolution remote sensing images." In: *Proceedings of IEEE International Congress on Image and Signal Processing*. CISP2010, 16-18 Oct., Yantai, China, pp. 2245-2250.

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

Maxwell, T. (2005). "Object-Oriented Classification: Classification of Pan-Sharpened QuickBird Imagery and a Fuzzy Approach to Improving Image Segmentation

Efficiency." M.Sc.E. thesis, Department of Geodesy and Geomatics Engineering Technical report No. 233, University of New Brunswick, Fredericton, New Brunswick, 157 pp.

Mayer, H. (1999). "Automatic Object Extraction from Aerial Imagery- A survey Focusing on Buildings." *Computer Vision and Image Understanding*, Vol. 74, No. 2, pp. 138-149.

Mayunga, S.D., D. J. Coleman and Y. Zhang (2007). "A semi-automated approach for extracting buildings from QuickBird imagery applied to informal settlement mapping." *International Journal of Remote Sensing*, Vol. 28, No. 10, pp. 2343-2357.

Möller, M., L. Lymburner, and M. Volk (2007). "The comparison index: A tool for assessing the accuracy of image segmentation." *International Journal of Applied Earth Observation and Geoinformation*, Vol. 9, No. 3, pp. 311–321.

Muller, S. and D.W. Zaum (2005). "Robust Building Detection in Aerial Images." In Proceedings of the Joint Workshop of ISPRS and the German Association for Pattern Recognition, 'Object Extraction for 3D City Models, Road Databases and Traffic Monitoring - Concepts, Algorithms, and Evaluation', Ed. Stilla U., F. Rottensteiner, and S. Hinz. CMRT05, IAPRS, Vol. 36, Part 3/W24, 29-30 Aug., Vienna, Austria.

Pal, N. R. and S. K. Pal (1993). "A review on image segmentation techniques." *Pattern Recognition*, Vol. 26, No. 9, pp. 1277–1294.

Peng J., D. Zhang and Y. Liu (2005). "An improved snake model for building detection from urban aerial images." *Pattern Recognition Letters*, Vol. 26, No. 5, pp. 587–595.

Richards, J. A. and X. Jia, (2006). *Remote Sensing Digital Image Analysis: An Introduction*. 4th ed., Springer-Verlag, Berlin Heidelberg, Germany.

Rottensteiner, F., J. Trinder, S. Clode, and K. Kubik (2003). "Building detection using LIDAR Data and multi-spectral images." In Proceedings of VIIth Digital Image Computing and Applications, Ed. Sun C., H. Talbot, S. Ourselin and T. Adriaansen. DICTA 2003, 10-12 Dec. Sydney, Australia, pp. 673-682.

Shackelford, A.K. and C.H. Davis (2003). "A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 9, pp. 1920-1932.

Shankar, B. U. (2007). "Novel classification and segmentation techniques with application to remotely sensed images." In Vol. 4400, *Lecture Notes on Computer*

Science: Transactions on Rough Set VII, Ed. J. F. Peters, A. Skowron, V. W. Marek, E. Orlowska, R. Slowinski, and W. Ziarko. Springer-Verlag, Berlin, Germany, pp. 295-380.

Sirmaçek, B., and C. Ünsalan (2008). "Building detection from aerial images using invariant color features and shadow information." In *IEEE International Symposium on Computer and Information Sciences*. ISCIS '08, 27-29 Oct., Istanbul, Turkey, pp. 1-5.

Song, Z., C. Pan, and Q. Yang (2006). "A region-based approach to building detection in densely build up high resolution satellite image." In: *Proceedings of IEEE International conference on Image Processing*. ICIP '06, 8-11 Oct., Taipei, Taiwan, pp. 3225-3228.

Stoney, W. E. (2008). " ASPRS Guide to land imaging satellites." [Internet] cited 12 Aug. 2011. Available in .pdf format from www.asprs.org/a/news/satellites/ASPRS_DATABASE_021208.pdf

Tian, J., and D. M. Chen (2007). "Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition." *International Journal of Remote Sensing,* Vol. 28, No. 20, pp. 4625-4644.

Tizhoosh, H.R. and H. Haussecker (2000), Fuzzy image processing: An overview. In *Computer Vision And Applications: A Guide for Students and Practitioners*, Ed. Jähne B. and H. Haußecker, Academic Press, Boston, USA, pp. 541-576.

Woodcock, C. E., and A. H. Strahler (1987). "The factor of scale in remote sensing." *Remote Sensing of Environment*. Vol. 21, No. 3, pp. 311-332.

Wuest, Ben A. (2008). Towards Improving Segmentation of Very High Resolution Satellite Imagery. M.Sc.E. thesis, Department of Geodesy and Geomatics Engineering Technical Report No. 261, University of New Brunswick, Fredericton, New Brunswick, Canada, 123 pp.

Zhang, Y. (2004). "Highlight Article: Understanding Image Fusion." *Photogrammetric Engineering & Remote Sensing,* Vol. 70, No. 6, pp. 675-661.

Zhang, Y., T. Maxwell, H. Tong, and V. Dey, (2010). "Development of a supervised software tool for automated determination of optimal segmentation parameters for eCognition." *Proceedings of ISPRS TC VII Symposium- 100 Years ISPRS*, 5-7 July, Vienna, Austria, pp. 690-696.

Zhou, W., A. Troy, and M. Grove (2008). "Object-based land cover classification and change analysis in the Baltimore metropolitan area using multi-temporal high resolution remote sensing data." *Sensors*, Vol. 8, No. 3, pp/ 1613-1636.

CHAPTER 2

IMAGE SEGMENTATION TECHNIQUES FOR URBAN LAND COVER SEGMENTATION OF VHR IMAGES: RECENT DEVELOPMENTS AND FUTURE PROSPECTS

This chapter contains a journal paper which has been submitted to an international journal and referred as:

Dey, V., Y. Zhang, M. Zhong, and B. Salehi (2011). "Image Segmentation Techniques for Urban Land Cover Segmentation of VHR Images: Recent Developments and Future Prospects." *International Journal of Remote Sensing,* (under review).

An abridged version of the presented journal paper is already published as a full paper-reviewed conference paper:

Dey, V., Y. Zhang, and M. Zhong (2010). "A Review of image segmentation techniques with remote sensing perspective". Wagner W., and B. Székely (Eds.). *Proceedings of ISPRS TC VII Symposium- 100 Years ISPRS*, 5-7 July, Vienna, Austria, pp. 31-42.

The paper presents a review of the current segmentation techniques used for urban land cover segmentation of VHR images. The paper identifies the advantages and disadvantages of each technique and concludes by identifying the most popular techniques. In order to present clearly, the original paper has been slightly edited.

Abstract

Due to continuous improvement of remote sensing sensor technology, spatial resolutions of satellite images have been refined to sub-metre resolution. Consequently, remote sensing image processing shifted from pixel-based to object-based image analysis (OBIA). OBIA follows closely with human perception of objects. The basic step of object formation in OBIA is image segmentation. Fortunately, image segmentation has been researched in computer vision for the last four decades. However, this doesn't alleviate the segmentation problem because image segmentation is domain specific. This paper reviews image segmentation techniques in the domain of urban land cover segmentation of very high spatial resolution (VHR) satellite images. The authors categorise the segmentation techniques into eight categories: clustering, level-set, Markov random field, fuzzy logic, neural network, multi-scale, watershed, and hierarchical split and merge (HSMR). The authors also describe recently developed techniques, deduce trends, (e.g., widely used techniques and commercially developed techniques) and elaborate on the potential techniques, where a developer can dig in for new developments. It is concluded that while all the categories have capability of efficient feature extraction, multi-scale and watershed stand ahead in the case of urban land cover segmentation based classification.

2.1 Introduction

Since the launch of the IKONOS satellite in 1999, remote sensing (RS) satellite sensor technology has been steadily developing. The spatial detail of satellite images has been improved to sub-metre resolution, e.g., the spatial resolution of Panchromatic (Pan) band of a GeoEye-1 satellite image is 0.5m. In addition, the traditional pixel-based analysis failed to handle the increased spatial variability within the land cover classes of VHR images [Blaschke and Strobl, 2001]. This failure and the demand of better results from VHR images led to the surge of object-based image analysis (OBIA) [Carleer et al., 2005; Blaschke et al., 2006; Blaschke, 2010].

OBIA, also known as geographic OBIA (GEOBIA) in RS, follows the logic of the human-based image interpretation [Hay and Castilla, 2006; Blaschke et al., 2006]. The basic step of OBIA is the generation of image objects using image segmentation [Blaschke, 2010]. Image objects are groups of pixels representing geographic classes such as buildings, trees, and grasslands [Blaschke et al., 2006; Hay and Castilla, 2006]. Fortunately, image segmentation has been widely studied in the field of computer vision and other domains (e.g., medical image, industry image, and range image) leading to hundreds of image segmentation techniques [Haralick and Shapiro, 1985; Reed and Buf, 1993; Pal and Pal, 1993; Cheng et al., 2001; Freixnet et al., 2002]. However, the techniques cannot be directly imported to RS because the choice of the image segmentation techniques is domain specific [Pal and Pal, 1993; Zouagui et al., 2004; Xia and Feng, 2009].

2.1.1 Description of Segmentation

For better understanding of the domain dependency, a general description of image segmentation is required. In general, image segmentation is defined as the process of completely partitioning an image into non-overlapping group of similar pixels called regions such that adjacent regions are heterogeneous [Pal and Pal, 1993]. Image segmentation is achieved based on two distinct properties of image intensity values: discontinuity (e.g., edge-based) and similarity (e.g., region-based) [Gonzalez and Woods, 2002]. Thus, the differences among various image segmentation techniques depend on the two major criteria: a) how the discontinuities or similarities (also known as homogeneity or heterogeneity measures) are evaluated; and b) how the pixels are aggregated (e.g., edge contour based and region based) [Gonzalez and Woods, 2002].

The above-mentioned criteria are domain specific. For example, in VHR RS domain, a multi-scale analysis is preferable because different ground objects need different intrinsic scales whereas, in medical image segmentation, the purpose of multi-scale processing is to reduce computational complexity [Pham et al., 2000; Gonzalez and Woods, 2002; Hay et al., 2003; Benz et al., 2004]. Hence, the domain of application has a great impact on the selection of image segmentation techniques.

2.1.2 Advantages of Segmentation based RS Image Analysis

Though image segmentation is often categorized as one of the most critical tasks in image processing, its benefits supersede its drawbacks in the RS domain [Pal and Pal, 1993; Blaschke et al., 2006]. The major benefits of image segmentation based object formation in the RS image analysis are as follows: a) it identifies image objects (regions) as perceived by a human eye; b) it enables the use of shape, size, and contextual information for analysis; c) it allows the use of topological relationships for vector based GIS operations; d) it decreases the execution time of classification and increases its accuracy ; e) it minimizes the modifiable areal unit problem (MAUP), caused by dependency of statistical results (e.g., mean and standard deviation) on the spatial units (the chosen spatial resolution of study); and f) it reduces the fuzzy boundary problems [Hay et al., 2003; Blaschke et al., 2006; Blaschke, 2010]. However, these benefits are the outcomes of an appropriate image segmentation technique of RS domain.

2.1.3 Segmentation: Dependent on Application

Similar to other domains, there are numerous different applications within RS domain and each application has different goals to achieve. Even in the same domain, the choice of image segmentation techniques depends on the goals of the application. For example, if the goal of segmentation based application in Figure 2.1 is to separate water and vegetation areas as well as lakes and rivers. Then, the deciding factors are: a) spectral properties for vegetation and water; and b) shape and size property for lake and river. Therefore, the required image segmentation techniques should utilize spectral as well as shape property of the image. However, if one changes the goal to the separation of water body and vegetation area, only the spectral property would suffice. The above example demonstrates how a simple change in the goal of the feature extraction application modified the choice. Hence, in complex applications (e.g., classification, change detection, impervious surface extraction, and building detection), a more involved analysis is required for the appropriate selection of the image segmentation technique [Shackelford and Davis, 2003; Falkowski et al., 2009; Werff and Meer, 2008; Zhou and Wang, 2008].

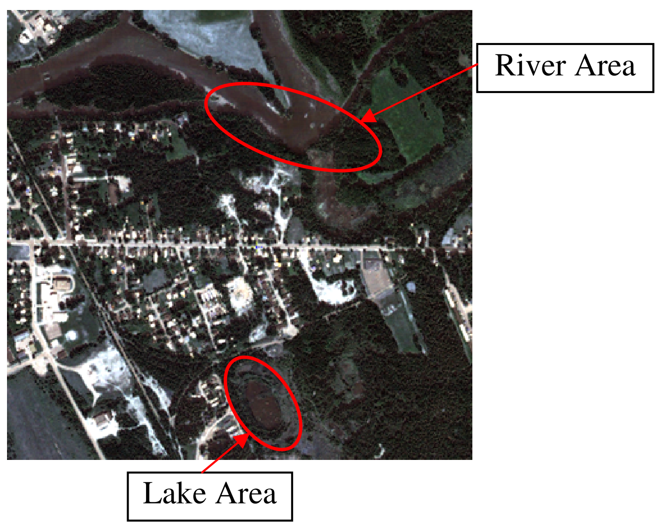


Figure 2.1: Example of spectral, shape, and size information needed for separating vegetation (green) and water (black) as well as river (linear) and lake (closed shape) from an image of Fredericton, Canada. (Source: Author's own image, © Quickbird)

2.1.4 Earlier Reviews on Image Segmentation

Image segmentation has long been studied in domains other than remote sensing [Fu and Mui, 1981; Haralick and Shapiro, 1985; Pal and Pal, 1993; Cheng et al., 2001]. Over the last decade, segmentation has been widely utilized in remote sensing [Blaschke et al., 2006; Blaschke, 2010]. The major review papers in the RS domain are Schiewe [2002], Carleer et al. [2005], Shankar [2007], and Blaschke [2010]. The first two review papers concentrated on with very few early methods in RS. The third review paper is relatively new and discussed the image segmentation techniques being used for RS in general. The final paper reviewed the development of OBIA applications, e.g., change detection, disaster and risk management, forest classification, and urban feature extraction. The final paper also mentioned the wide use of multiresolution segmentation, which is implemented in the commercial software eCognition™, for the OBIA applications [Baatz and Schäpe, 2000].

In addition, it is also important to identify the general rules on how to choose the image segmentation techniques for RS specific applications (e.g., change detection, urban feature detection and classification, and forest stand classification), which are not focused in the earlier reviews. Further, it is also important to identify the widely-used or potential techniques for VHR images, apart from multiresolution segmentation mentioned by Blaschke [2010]. This review paper attempts to fill the gaps in these two areas.

2.1.5 Contribution of This Paper

Urban land cover is complex and has high frequency of significant image object changes [Herold et al., 2003; Mesev, 2003]. This necessitates both an accurate and frequent information extraction of urban land covers [Mesev, 2003]. Hence, this paper identified eight major categories of urban land cover segmentation techniques from earlier reviews [Carleer et al., 2005, Shankar 2007; Blaschke, 2010]: clustering, level-set, Markov random field (MRF), artificial neural network (ANN), fuzzy logic based, multi-scale (MSc), watershed, and hierarchical split and merge (HSMR). A brief introduction of each category is provided followed by the concise description of their recently developed techniques.

The paper also identifies the suitability of each technique in terms of the applications, broadly classified as feature extraction and classification. Finally, the paper concludes with discussions of the following: a) widely-used urban land cover VHR image segmentation techniques; b) image properties of importance (e.g., spatial, contextual, and prior knowledge); c) user-defined parameters complexity; d) commercial implementation; and e) accuracy assessment of segmentation.

The rest of the paper is organized as follows. Section 2 discusses the rules/factors affecting the choices of segmentation techniques. Section 3 elaborates on the various existing techniques for RS image segmentation, and their advantages and disadvantages in achieving urban land cover segmentation. Section 4 specifies conclusions, based on the review of the eight categories, related to widely-used techniques, the parameter complexity, the techniques implemented in the form of software, accuracy assessments, and the techniques with the potential for future developments.

2.2 The Factors Governing the Choices of Segmentation Techniques

A human interpreter is the best interpreter of a VHR image [Lang et al., 2009]. A human image interpreter uses visible cues for image analysis, which are synonymous to image interpretation elements [Estes, 1999]. Hence, these image interpretation elements are among the most important factors for the selection of a suitable segmentation technique for an application. Other important factors comprise the performance of segmentation techniques and ease in usage of the technique for a segmentation based application.

2.2.1 Image Interpretation Elements

The image interpretation elements are basically the properties of pixel values, either viewed alone or in a group (e.g. in spatial domain). These image interpretation elements or image properties used for image segmentation are as follows: a) spectral, b) spatial, c) texture, d) shape, e) size, f) context, g) shadow, h) connectivity, and i) association.

Connectivity, shadow, and association can be categorised as part of prior knowledge about an image [Estes, 1999; Baltsavias, 2004]. In general, prior knowledge is defined as human knowledge about the behaviour and pattern of the ground objects, e.g. the number of land cover classes in the image scene, ancillary data (GIS shape layers), and probability distribution [Baltsavias, 2004].

All of the existing image segmentation techniques utilize one or more of these interpretation elements. For example, Hu et al. [2005] and Triaz-Sanz et al. [2008] used spectral, scale, and several texture based features; Baltsavias [2004] reviewed several techniques and identified how knowledge has been integrated into those techniques; and Baatz and Schäpe [2000] used standard deviation of spectral values, shape, and size based features. Figure 2.1 demonstrates that how the change of interpretation elements changes the segmentation technique. The Figure also demonstrate that different elements are favoured depending upon the goal of the application.

Both spectral and spatial property are the most important interpretation elements because image segmentation requires aggregation of pixels in the spatial domain [Pal and Pal, 1993; Hay et al., 2003; Blaschke et al., 2006]. Since all the techniques utilize spectral property, it can be assumed that the techniques which lack the utilization of spatial property are ineffective for urban land cover segmentation of a VHR image [Hay et al., 2003; Benz et al., 2004; Blaschke, 2010].

2.2.2 Complexity in Usage

Given the wide variety of segmentation algorithms, it is possible to encounter several techniques, which utilize the same interpretation elements. Hence, the obvious choice would be to choose the technique with the least complexity related to the parameters as well as usage. The complexity in usage primarily refers to whether usable software exists for a particular technique or not. In general, commercially available software are considered to be usable. If usable software is not available, it may take months to develop the usable software from scratch. This is because the software has to go through several phases of software development even after programming of the codes (or implementation), e.g., software testing and documenting [Mishra and Zhang, 2011]. Apart from the software, a large number of user-defined parameters, which have high impact on nature of the segmentation results, also increase the complexity in usage of the segmentation technique.

Apart from software availability, the customization ability of software is also significant for its popularity. For example, popular segmentation software eCognition™ Developer 8 provide the ability using customized features and rule sets (now owned by Trimble Inc.); and ENVI™ (owned by ITT Visual Information solutions) provides the ability with the interactive data language (IDL) for easy visualization and programming [Definiens AG, 2009; ITT Visual Information Solutions, 2011].

2.2.3 Assessment Factor

The next factor is the assessment of the image segmentation results. Marpu et al. [2010] compared twelve different segmentation techniques of different commercial software based on their proposed quantitative assessment technique. They found that the results of multiresolution segmentation of Definiens Developer (now eCognition™ Developer 8) are among the best. Such a comparison helps in identifying the best segmentation technique for a specific application (e.g., building detection and land cover classification). Although several quantitative assessment techniques exist, the notion of visually pleasing segmentation results is still widely popular [Zhang, 1997; Zhang et al., 2008; Neubert et al., 2008, Lang et al., 2009; Marpu et al., 2010; Corcoran et al., 2010].

All the factors stated in the above sections have been categorized into three different groups: 1) concept based factors, 2) usage based factors, and 3) assessment/evaluation based factors. Image interpretation elements are major constituents of concept based factors. Concept based factors also include top-down and bottom-up approaches (explained in the Section 2.3.1) of segmentation, supervised and unsupervised approach, and scale factor of segmentation [Guindon, 1997; Wuest and Zhang, 2009; Corcoran et al., 2010].

On the other hand, usage based factors consist of availability of usable software and the number of user-defined parameters. Similarly, evaluation based factor consist of the nature of evaluation approach, i.e., quantitative evaluation or visual assessment of the segmentation results. Table-2.1 summarizes these factors for selecting image segmentation techniques. The next Section describes a total of eleven recently developed and commonly used image segmentation techniques in the view of the above-mentioned factors.

Table 2.3: The rules/factors governing the choices of selection of segmentation techniques for remote sensing.

<table>
<tr><th>Conceptual factors</th><th>Usage factors</th><th>Evaluation factor</th></tr>
<tr><td>Top-down and bottom-up approach</td><td rowspan="4">Number of user-defined parameters and implementation complexity, i.e.., existence of software and its ease in usage and customization</td><td>Quantitative evaluation, e.g., Zhang [1997], Corcoran et al. [2010]</td></tr>
<tr><td>Supervised and Unsupervised approach</td><td rowspan="3">Visual assessment</td></tr>
<tr><td>Multi-scale or hierarchical</td></tr>
<tr><td>Interpretation elements, e.g., Spectral, spatial, contextual, shape, size and prior knowledge</td></tr>
</table>

2.3 Review of Recent Urban Land Cover Segmentation Techniques

2.3.1 Traditional Categorization Schemes

With hundreds of image segmentation techniques in place, it is necessary to categorize them for proper representation. The traditional categorization scheme of image segmentation has four categories: a) pixel/point/threshold based, b) edge based, c) region based, and d) hybrid [Spirkovska, 1993; Schiewe, 2002]. Pixel-based techniques utilizes global threshold generally derived from image histograms. These thresholds are used to evaluate the similarity of pixels. Based on the similarity, these pixels are aggregated to form the regions. On the other hand, edge-based techniques detects the boundaries of regions and then close the boundaries to form regions [Pal and Pal, 1993; Schiewe, 2002; Blaschke et al., 2006].

Region based techniques are divided into region growing, merging and splitting, and their combinations [Blaschke et al., 2006]. Region growing (also a bottom up approach) starts from a single seed pixel region. This region is grown by including neighbouring pixels until a homogeneity/heterogeneity criterion for the inclusion is satisfied. For example, multiresolution segmentation uses region growing for segmentation [Benz et al., 2004]. On the other hand, region merging and splitting (also a top-down approach) starts by splitting the image into sub-regions and later these regions are merged based on a heterogeneity/homogeneity criterion, e.g., hierarchical split and merge technique (HSMR) [Ojala and Pietikäinen, 1999]. A hybrid technique includes a fusion of one or more of pixel based, edge based, or region based. Traditional categorization scheme excludes the indication of image interpretation elements used for segmentation.

2.3.2 Model Based Categorization

Image segmentation techniques can also be categorized based on the used major image interpretation elements, e.g., texture-based by Reed and Buf [1993]; and color and texture-based by Guo et al. [2005]. However, there are numerous techniques which utilize one or more interpretation elements [Shankar, 2007]. Hence, to be more specific to the description of techniques, a model/approach-based categorization scheme is selected [Shankar, 2007]. The models/approaches provide explicit information regarding steps of the techniques, used interpretation elements, and possible modifications in the approaches of techniques. The models/approaches selected for the review are as follows: a) Clustering approach, b) Level-set, c) MRF model, d) ANN model, e) Fuzzy model, f) MSc model, g) Watershed model, and h) HSMR model.

The models stated above can be further categorized as: i) mathematical models namely, probability and statistics based optimization model (Level set, MRF and ANN model), and fuzzy logic based model; and ii) conceptual model (MSc, Watershed and HSMR model). It is important to note that the above mentioned models/approaches are by no means complete in categorizing all the RS segmentation techniques but they do represent the major ones for VHR urban land cover segmentation [Carleer et al., 2005; Shankar et al., 2007].

Few other model/approaches namely, object-background model and edge-based approach are obsolete. Hence, these two models are not discussed in here. Interested readers can go to Pal and Pal [1993] for the details of these models/techniques. The next few sub-Sections describe the recently developed urban land cover segmentation techniques, their interpretation elements, the parameters and usage complexity, and images of application.

2.3.3 Clustering Approaches

Clustering is based on the concept of pixel grouping but it is conceptually different from segmentation. While traditional clustering techniques (K-means and ISODATA) relies on aggregation in spectral measurement space, segmentation relies on aggregation in the spatial domain [Haralick and Shapiro, 1985]. Haralick and Shapiro [1985] provided a good review of clustering-based segmentation techniques with spatial domain and called it as spatial clustering. Even with the spatial domain, most of the clustering techniques need initial number of clusters (segments), which is difficult to estimate for unsupervised segmentation [Pal and Pal, 1993]. Hence, a successful clustering-based segmentation technique for VHR image needs inclusion of spatial domain and automatic determination of initial number of cluster/segments.

Wang et al. [2010] proposed a region based image segmentation (RISA), which is a hybrid of K-means clustering and region merging approaches. RISA utilises spectral, spatial, shape, and size properties. However, the technique suffers from parameter complexity because it requires more than five user-defined parameters to estimate.

However, the process has been successfully implemented as software and has been applied to an urban Quickbird image with the results reportedly comparable to that of the eCognition's multiresolution segmentation technique (a very popular technique, as mentioned by Blaschke [2010]). Further, the technique offers multi-scale analysis, which is crucial in VHR segmentation [Hay et al., 2003; Benz et al., 2004; Blaschke et al., 2006].

In general, clustering based techniques are unpopular for VHR land cover image segmentation because of the above-mentioned conditions. Hence, the authors have not included any research papers based on clustering. Moreover, the authors do not suggest clustering based segmentation for an urban VHR image.

2.3.4 Mathematical Models Based Image Segmentation

2.3.4.1 Level Set Model

Level set model (also formulated as active contour model or snake model) tracks boundaries of the object by minimizing the defined energy function with appropriate boundary conditions [Peng et al., 2005; Karantzalos and Argialas, 2009]. Level-set model has been recently used in urban remote sensing for segmentation applications, e.g., urban feature extraction, buildings, and roads, by Karantzalos and Argialas [2009], and urban change detection by Bazi and Melgani [2010]. The authors found very few relevant research papers based on level set models VHR land cover segmentation. The relevance of the papers is determined on the basis of the factors stated in the Section 2.2. Hence, the authors have not included any segmentation techniques based on Level set and recommend for more experimentation in this field.

2.3.4.2 MRF Model

While clustering performs optimization in measurement space, MRF is based on statistical and probabilistic theory based optimization. Image segmentation problem in MRF is represented as discrete labelling problem. The objective function of the labelling problem is generally formulated using probabilistic estimation, e.g., maximizing a posterior (MAP) estimation, and maximizing posterior marginal (MPM) [Li, 2009]. However, direct solution of the MAP/MPM estimation of MRF is impossible because of the mathematical complexity of its probability based estimation.

However, due to equivalence of MRF with Gibbs random field (GRF) [Geman and Geman, 1984], probability functions of MRF can be representing using energy potentials. These energy potentials require neighbourhood based interaction of pixels, as shown in Figure 2.2. Hence, MRF models employ spatial context using potentials interactions (mathematically known as Markovian property) and prior knowledge (of probability distributions). Subsequently, MRF represents an attractive technique for texture analysis and spatial segmentation [Bouman and Shapiro, 1994; Poggi et al., 2005].

P(1,1)	P(1,2)	P(1,3)	P(1,4)	P(1,5)
P(2,1)	P(2,2)	P(2,3)	P(2,4)	P(2,5)
P(3,1)	P(3,2)	P(3,3)	P(3,4)	P(3,5)
P(4,1)	P(4,2)	P(4,3)	P(4,4)	P(4,5)
P(5,1)	P(5,2)	P(5,3)	P(5,4)	P(5,5)

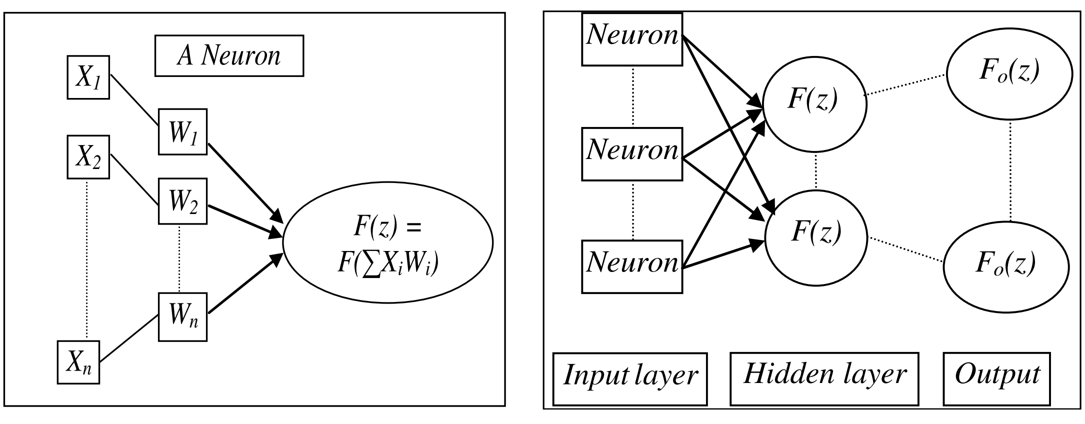
Figure 2.2: P(3,3) represents pixel and its 8 point neighbourhood in gray.

Although MRF uses spatial context, MRF based techniques have not been implemented in commercial software. Moreover, MRF requires prior knowledge of an initial number of segment labels [D'Elia et al., 2003; Poggi et al., 2005]. In addition, the classes of an urban land cover VHR image are too complex to be modeled efficiently by statistical distributions, which are required for MRF based segmentation [Herold et al., 2003; Platt and Rapoza, 2008]. Due to the above-mentioned disadvantages, the authors have not included and do not recommend MRF based segmentation techniques for VHR image segmentation

2.3.4.3 ANN model

While MRF is probabilistic optimization, ANN is a machine learning based optimization technique. ANN simulates the functioning of the human brain processing element, i.e., neurons [Tso and Mather, 2004]. These neurons are cobwebbed to form a learning network which requires training data and produces a generalized framework for segmentation/classification of rest of the data, as shown in Figure 2.3 [a] and [b]. A conceptual and detailed description of ANN with respect to remote sensing can be found in Tso and Mather [2004], Mather [2004], Atkinson and Tatnall [1997], and Mas and Flores [2008]. Each of them identified major advantages and disadvantages of ANN in general.

As identified by them, the major advantage of ANN a based segmentation technique is that it does not require any statistical distributions. However, the design of networks of ANN is very complex and requires experience of the operator to achieve an effective design [Mather, 2004; Mas and Flores, 2008]. Due to these disadvantages, the authors do not suggest ANN based segmentation techniques for land cover segmentation of an urban VHR image.



[a] Structure of a Neuron [b] Structure of a Neural Network

Figure 2.3: [*a*] shows structure of a simple neuron with inputs as X_i s, linearly combined with weights W_i s to form *z,* and then *z* is passed on to a threshold function to get the output. [*b*] shows formation of a simple neural network by combination of several neurons with hidden layers represented as dots with different threshold functions, i.e., F & F_o at different layers.

2.3.4.4 Fuzzy Model

In a VHR image, ambiguity/fuzziness across segment boundaries is unavoidable (Benz et al., 2004). This makes fuzzy logic a better candidate compared to MRF and ANN models. Further, fuzzy logic has been recently used in urban VHR based applications, such as classification, feature extraction, and change detection [Benz et al., 2004; Hester et al., 2010; Mohammadzadeh and Zoej, 2010; Aldred and Wang, 2011]. A conceptual and detailed description of general methodology of fuzzy logic applicable to image segmentation can be found in Tso and Mather (2004) and Tizhoosh and Haussecker (2000).

The most researched fuzzy based segmentation technique is Fuzzy-C-Means (FCM) based clustering. Similar to traditional clustering, the basic FCM requires an initial number of clusters and lacks utilization of spatial domain [Fan et al., 2009]. Although several techniques have been proposed for minimizing the above-mentioned problems, none of them is widely-used or commercially implemented in software [Fan et al., 2009; Hasanzadeh and Kasaei, 2010].

Apart from FCM, Fuzzy Image Regions Method (FIRME), proposed by Lizarazo and Barros [2010], is among recently developed technique. However, the technique is focused on fuzzy segmentation based classification as an integrated process. Hence, the technique is unsuitable for segmentation based applications other than the classification.

Overall, it can be concluded that fuzzy segmentation too is not widely popular for general VHR image segmentation. However, fuzzy based techniques have high customization possibilities, which are suitable for supervised segmentation [Mohammadzadeh and Zoej, 2010; Aldred and Wang, 2011].

2.3.5 Conceptual Models Based Image Segmentation

The conceptual models (MSc model, watershed, and HSMR model) of a VHR image refer to how the image objects can be best represented and analyzed for effective segmentation. For example, Figure 2.4 shows that the segmentation should be multi-scale for effective identification of both objects. The next few sub-Sections describe the recently developed and widely used techniques of the above mentioned three conceptual models namely, MSc, watershed, and HSMR.

2.3.5.1 Multi-scale Model

It has been long established that scale is important in the analysis of RS imagery [Woodcock and Strahler, 1987]. Further, a single scale is often considered inappropriate for the analysis which should be carried out at different hierarchical scales for meaningful object extraction [Benz et al., 2004; Ju et al., 2005; Platt and Rapoza, 2008]. Scale of a meaningful object can be defined as the level of aggregation and abstraction at which an object can be best described [Benz et al., 2004]. Hence, an object which is smaller than the spatial resolution of image cannot be identified because of the scale of observation is limited by the resolution, as shown in Figure 2.4.

The concept of meaningful object varies based on the applications and the image interpreter. For example, in feature extraction, each building is a meaningful object and improper segmentation of road is not a problem whereas, in segmentation based classification, both roads and buildings constitute meaningful objects [Benz et al., 2004; McGuinness and O'Connor, 2010]. Hence, multi-scale analysis is often desired in RS image segmentation because the analysis aims to extract meaningful objects.

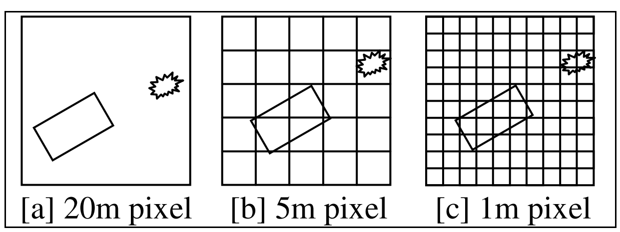


Figure 2.4: Shows the concept of the appropriate scale of representation of an object. Two objects of rectangular, R, and star, S, shaped have been taken as examples, where [a] ,at 20m spatial resolution (SR), shows that both objects are undetectable; [b], at 5m SR, shows that R is detectable but S is not and [c], at 1.25m SR, shows that both are detectable. (Image Source: Blaschke [2010], Courtesy: ISPRS, Elsevier)

2.3.5.1.1 Selected techniques

The most widely-used technique of MSc model is multiresolution segmentation, which is based on multi-fractal analysis known as fractal net evaluation approach (FNEA) [Baatz and Schäpe, 2000; Blaschke et al., 2006; Blaschke, 2010]. Multiresolution segmentation has been implemented in commercial software called as eCognition™ Developer (now owned by Trimble) since 2000 [Blaschke, 2010]. Moreover, Neubert et al. [2008] and Marpu et al. [2010] concluded by quantitative evaluation of twelve segmentation based software that eCognition™ has one of the best segmentation technique for urban land cover segmentation as well as other RS applications.

Although widely popular, FNEA needs an effective estimation of three user-defined parameters (scale value, shape weight, and compactness weight) for appropriate segmentation [Hay et al., 2003; Tian and Chen, 2007]. The parameter estimation is also a major problem of multiresolution segmentation. Several researchers have proposed approaches to estimate the effective parameters, e.g., Maxwell and Zhang [2005], Tian and Chen [2007], Costa et al. [2008], Drăgut et al. [2010], However, wide applicability of these approaches is still to be established.

Hay et al. [2003] and Hay and Marceau [2004] described two more MSc based techniques namely, Linear scale-space and blob-feature detection (SS), and multiscale object-specific analysis (MOSA). Both MOSA and SS are based on scale space theory

[Lindberg, 1994]. As per Hay et al. [2003], both MOSA and SS have no parameters complexity and are conceptually sound. However, they are not widely popular because of lack of implementation in software format as well as lack of comprehensive testing on a wide variety of VHR images.

To summarize, MSc model based techniques are among widely-used and recognized techniques for urban land cover segmentation [Hay et al., 2003; Blaschke, 2010]. Among MSc model based techniques, multiresolution segmentation is the most popular technique. Further, most of the researcher ascertained that multiscale segmentation is the most effective way of urban land cover segmentation of VHR image [Hay et al., 2003; Blaschke et al., 2006; Platt and Rapoza, 2008, Blaschke, 2010]. Hence, the authors recommend the multiscale based segmentation techniques to be among the best for an urban VHR image segmentation.

2.3.5.2 Watershed Model

Watershed segmentation is another conceptual model for image segmentation. Watershed model views an image as a topographic surface. It assumes that if water effuses out from selected minimum points across the image, then the boundaries where the flooded regions from each minimum point meet constitute the desired segmentation regions, see Figure 2.5 [a], [b], [c] & [d] [Beucher, 1992]. Due to this representation, watershed segmentation is also known as a morphological segmentation technique

[Beucher, 1992; Pesaresi and Benediktsson, 2001]. One of the most effective implementation of watershed segmentation is marker controlled (MC) watershed segmentation, which was proposed by Meyer and Beucher [1990]. Hence, watershed based techniques are relatively new because they have been used for segmentation for the last two decades. On the other hand, MRF, ANN, and Fuzzy based techniques have been used for segmentation for more than two decades.

The major challenge of a watershed based technique is to reduce over-segmentation, which is the outcome of traditional watershed segmentation technique [Chen et al., 2006; Li and Xiao, 2007; Castilla et al., 2008]. The next few paragraphs describe three examples of recently developed watershed segmentation techniques combating the above mentioned challenges. These three examples are applied for urban VHR image segmentation. These three examples are chosen by the authors as the potential techniques for an effective segmentation.

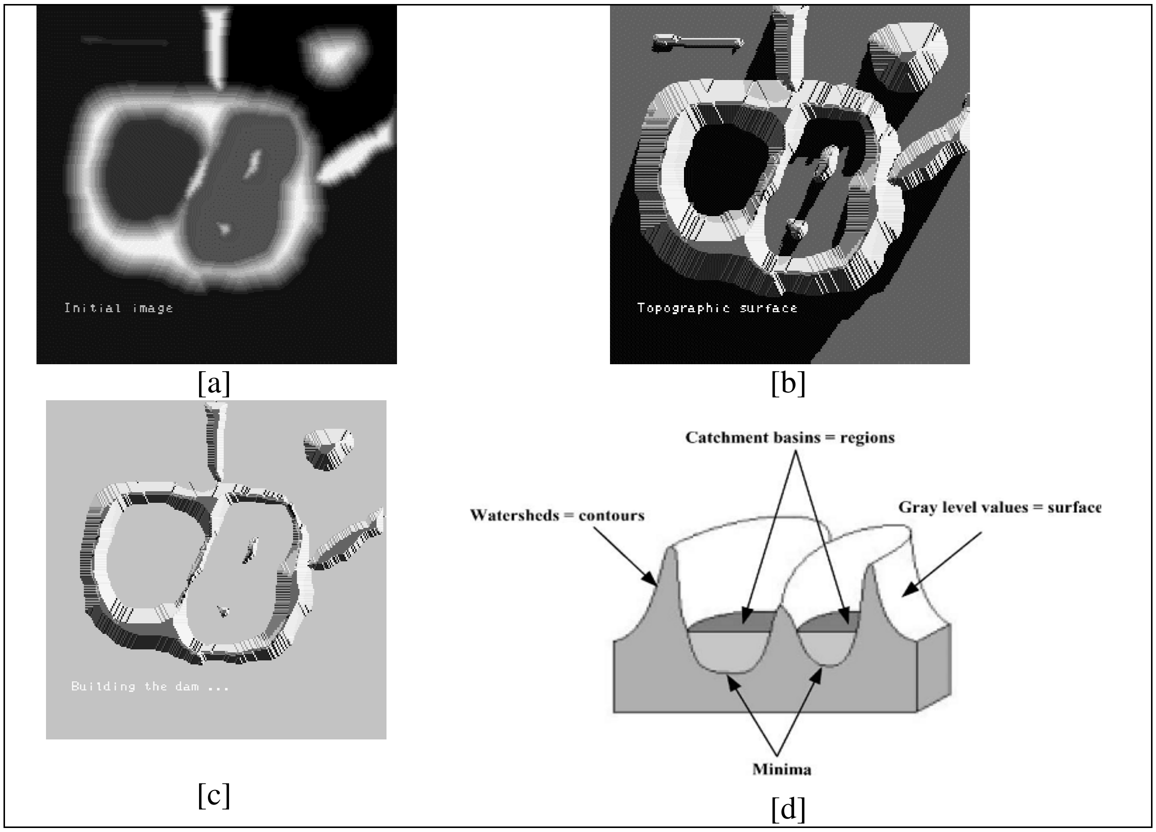


Figure 2.5: [a] shows a simple grey level medical image; [*b*] shows the topographical transformation of [a] by representing elevation as grey levels; [*c*] shows the flooding process from the minimum: the flooding process goes until flooding from two different minima meet; and [*d*] shows a cross section of WS catchment basins where the flooding from the minima meet. (Source [a, b & c]: http://cmm.ensmp.fr /~beucher/wtshed.html, Copyright: Beucher [2010]; Source [d]: Kim and Kim [2003], Courtesy: Pattern Recognition Letters, Elsevier.)

2.3.5.2.1 Selected techniques

MOSA, a multi-scale and modified MC based watershed segmentation, was proposed by Hay and Marceau [2004]. As mentioned in the Section 2.3.5.1.1, MOSA uses scale space theory and has complex mathematical representation. In spite of less user-defined parameters and strong conceptual representation of the segmentation problem, MOSA is not easy to use because of lack of implementation in software [Hay et al., 2003].

Another recently developed watershed based technique is size-constrained region merging (SCRM) proposed by Castilla et al. [2008]. They also used MC based watershed technique after smoothing of the image by the gradient inverse-weighted edge-preserving smoothing algorithm. The technique resulted in an over-segmented image. To merge the over-segmented image, the technique required three user-defined parameters: 1) the desired mean size of output segments (in hectares); 2) the minimum size required for segments (in hectares); and 3) desired spatial accuracy of boundaries (in metres). However, in the present state, the method lacks a multi-scale representation. Nevertheless, the technique has been implemented in software and showed a comparable performance with respect to multiresolution segmentation based on the experiment by Marpu et al. [2010].

Watershed based segmentation techniques followed by region merging has shown increasing interest among the researchers [Wang, 1997; Hay and Marceau, 2004; Castilla et al., 2008]. The main reasons for the interest are as follows: i) logical/conceptual representation of segmentation goal; ii) low parameter complexity; iii) capability of multi-scale representation; and iv) capability of utilization of spectral, spatial, shape, and texture with flexibility to include more [Wang, 1997; Hay and Marceau , 2004; Castilla et al., 2008; Wu et al., 2009]. All these factors suggest that watershed based segmentation techniques are potential techniques for urban VHR image segmentation.

2.3.5.3 HSMR model

HSMR is a top-down approach (split and merge) based conceptual model of segmentation. HSMR model, proposed by Ojala and Pietikäinen [1999], has three basic steps: a) hierarchical splitting, b) agglomerative merging, and c) boundary refinement. Each of the steps has been shown in Figure 2.6[a], [b], and [c]. HSMR based techniques have four user-defined parameters. Hence, the techniques require an operator's experience for effective segmentation.

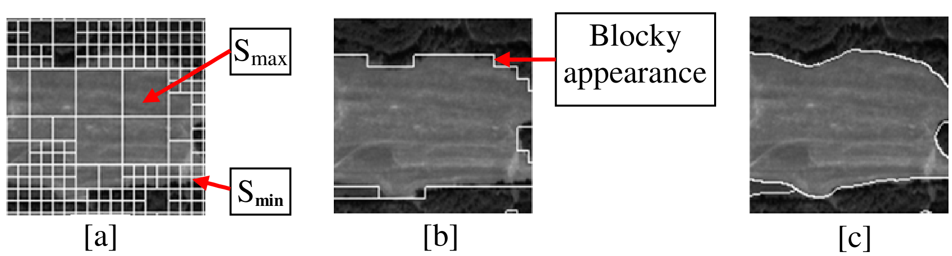


Figure 2.6: [*a*] shows hierarchical splitting with S_{max} and S_{min} as sizes of maximum and minimum blocks; [*b*] shows the blocky appearance of boundaries generated after agglomerative merging step is performed on split regions; [*c*] shows smoothened boundaries after boundary refinement of the image produced at step [*b*]. Source: Wuest and Zhang [2009]; Courtesy: ISPRS, Elsevier.

Texture feature defined by Ojala and Pietikäinen [1999] was applicable only for grey level images. Hu et al. [2005] modified the HSMR technique and used it for segmentation of a multispectral Quickbird image. However, Wuest and Zhang [2009] found several discrepancies in the method of Hu et al. [2005] namely, fragmentation and discontinuous boundaries of regions. They proposed a fuzzy integration in the HSMR model for supervised image segmentation using five classes namely, forest, grass, soil, water and urban. But, the technique proposed by Wuest and Zhang [2009] is restricted to the specified five classes, which is a huge disadvantage. However, the technique can be extended to unsupervised segmentation as proposed by Wuest and Zhang [2008].

The major drawback of HSMR based segmentation technique is its estimation of four user-defined parameters [Hu et al., 2005]. Further, same set of texture features might not be appropriate for discrimination of objects across different images because of the complexity of urban classes (as described in the Section 2.1.5). However, HSMR model follows a top-down approach which may be useful in the identification of heterogeneous land use (as demonstrated by Wuest and Zhang [2009]). Overall, the authors do not recommend HSMR based techniques for segmentation of an urban VHR image.

2.4 Conclusions

2.4.1 The Summary

This study reviewed recently developed remote sensing image segmentation techniques related to land cover segmentation of urban VHR imagery. In order to justify the selection of techniques, several factors which govern the choice/popularity of a particular image segmentation technique were identified. These factors are:

1. Utilization of image interpretation elements: spectral, spatial, texture, shape, size, context, shadow, connectivity, and association;
2. Utilization of Multi-scale concept;
3. Estimation and number of parameters;
4. Usage complexity; and
5. Evaluation of segmentation.

Then, eight major categories related to techniques of urban land cover segmentation using VHR imagery were identified: clustering, level set, MRF, ANN, fuzzy, multi-scale, watershed, and HSMR. Basic concept of each of the categories is briefly explained along with their advantages and disadvantages. This was followed by a brief discussion of recently developed techniques and their potentials for land cover segmentation of an urban VHR image. Image interpretation elements, evaluation measures, and the software availability status of each of the techniques were also specified. It is important to note that the number of parameters and interpretation elements of a technique mentioned in this review paper are as specified by the research papers which proposed the techniques.

Finally, the potential of each technique towards land cover segmentation based classification and feature extraction from an urban VHR image has been justified. The next paragraph deals with the topic related to potential techniques in more detail.

2.4.2 Widely-used Techniques and their Applications

Regarding widely-used techniques, it was found that MRF, ANN, and fuzzy are not popular for urban VHR applications, e.g., feature extraction and classification. Most of the traditional clustering, MRF, and FCM based fuzzy techniques require an initial number of segments. This number is difficult to estimate for general unsupervised segmentation. Hence, these techniques are unsuitable for unsupervised segmentation. On the other hand, ANN based techniques require experience for effective segmentation. To summarize, the mathematical/probabilistic models (MRF, ANN, and Fuzzy) are still unable to represent the complex RS ground image in general.

Contrary to mathematical models, conceptually derived heuristics models (MSc and watershed) are more popular for segmentation of an urban VHR image. MSc based techniques are popular because of its effectiveness in analyzing a VHR image at appropriate scales [Hay et al., 2003]. Moreover, the techniques of all the eight categories employ multi-scale analysis. On the other hand, watershed based techniques are gaining popularity because of their customization abilities. The potential techniques in the category of watershed models are MOSA and SCRM whereas, in MSc model, FNEA is the most popular [Hay et al., 2003; Benz et al., 2004; Castilla et al., 2008].

Level set and HSMR are recently developed techniques. While Level set based techniques are more suited for feature extraction application instead of segmentation, HSMR based techniques require more experiments to deduce any trend regarding its application. However, the authors conclude that at present both HSMR and Level set based techniques are not recommended for segmentation of an urban VHR image in general. Table 2.2 and 2.3 summarize the features of the techniques reviewed in this paper under each category: (1) clustering; (2) Mathematical models (MRF, ANN, Fuzzy, and Level set); and (3) Conceptual models (Multi-scale, Watershed, HSMR).

Table 2.2: Enlists the selected segmentation techniques of clustering, MRF, ANN, Fuzzy, and Level Set models with their used image interpretation elements, accuracy assessment techniques, and applications.

Clustering approach					
Authors (method)	**Categorisation**		**Image used**	**Evaluation**	**Application**
	Interpretation elements	**Approach**		**Method**	
Wang et al. [2010b] (RISA)	Spectral, Spatial, Scale,Shape and Size	Region Growing & Merging	SPOT-5 & Quickbird	Classification accuracy	Urban Area (implemented as software)
Fuzzy model					
Fan et al [2009] (SWFCM)	Spectral, spatial & Prior Knowledge	Cluster growing	Landsat TM	Classification accuracy and cluster validity indices	Agriculture mixed water land
Hasanzadeh and Kasaei [2010]	Spectral and Spatial	Cluster growing	Landsat-7	Quantitative segmentation accuracy	Agricultural
Lizarazo and Barros [2010] (FIRME model)	Spectral , spatial and contextual	Region growing	Quickbird	Classification accuracy 83%	Urban
Level Set					
Karantzalos and Argialas [2009]	Spectral, shape size and scale	Region based	Quickbird	Visual assessment	Urban

Table 2.3: Enlists the selected segmentation techniques of clustering, multi-scale, watershed, and HSMR models with their used image interpretation elements, accuracy assessment techniques, and applications.

Authors (method)	Categorisation: Interpretation elements	Categorisation: Approach	Image used	Evaluation: Method	Application
Multi-scale model					
Baatz and Schäpe [2000]	Spectral, spatial, size and shape	MR segmentation	Almost all VHR RS imagery	Visual assessment	Implemented as Software
MOSA by Hay and Marceau [2004]	Spectral, size, scale and spatial	Region based	IKONOS Pan	Visual assessment	Forest
Blob feature detection (SS) By Hay et al. [2003]	Spectral, size, scale, connectivity and spatial	Region based	IKONOS Pan	Visual assessment	Forest
Watershed model					
SCRM by Castilla et al. [2008]	Spectral, shape, size and spatial	Region growing	Quickbird	Visual assessment	Agricultural and urban
HSMR model					
Hu et al [2005]	Spectral, texture, size and scale	Region splitting and Merging	Quickbird and IKONOS (MS and Pan)	Visual assessment	Urban
Wuest and Zhang [2008]	Spectral, texture, size and scale	Region splitting and Merging	Quickbird MS	Visual assessment	Urban

2.4.3 Evaluation Measures

The final discussion concerns evaluation measures used for image segmentation. Most of the evaluation measures are based on the under-segmentation or over-segmentation of a segment [Zhang et al., 1997; Corcoran et al., 2010]. In spite of considerable progress in evaluation measures, visual assessment is still widely used and required [Corcoran et al., 2010]. However, most of the quantitative measures are more of an individual segment based comparison measures, i.e., local segmentation quality evaluation.

The authors recommend that along with the experimentation and research on potential segmentation techniques, a research on an effective global segmentation quality assessment measure should be conducted. Moreover, the evaluation should be fast and easy to implement as well as analyze.

ACKNOWLEDGEMENTS

This research is funded by the Strategic Project Grants Program of the Natural Science and Engineering Research Council of Canada (NSERC). The authors would also like to thank reviewers for their suggestions in improving this paper.

REFERENCES

Aldred, D. A., and J. Wang (2011). "A method for obtaining and applying classification parameters in object-based urban rooftop extraction from VHR multispectral images." *International Journal of Remote Sensing*, Vol. 32, No. 10, pp. 2811-2823

Atkinson, P.M., and A. R. L. Tatnall (1997). "Introduction Neural networks in remote sensing." *International Journal of Remote Sensing,* Vol. 18, No. 4, pp. 699.

Baatz, M. and A. Schape (2000). "Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation." *Angewandte*

Geographische Informations-Verarbeitung XII. Ed. J. Strobl et al. AGIT Symposium, Salzburg, Germany, 2000, pp. 12-23. [Internet], cited 12 Aug. 2011. Available in .pdf format through www.ecognition.cc/download/baatz_schaepe.pdf.

Baltsavias, E.P. (2004). "Object extraction and revision by image analysis usingexisting geodata and knowledge: current status and steps towards operational systems." *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 58, No. 3-4, pp. 129-151.

Bazi, Y., and F. Melgani (2010). "Unsupervised change detection in multispectral remotely sensed imagery with level set models." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, No. 8, pp. 3178-3187.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I. and M. Heynen (2004). "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information." *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 58, No.3-4, pp. 239-258.

Beucher, S. (1992). "The watershed transformation applied to image segmentation." In *10th Pfefferkorn Conf. on Signal and Image Processing in Microscopy and Microanalysis,* 16-19 sept. Cambridge, UK*,* pp. 299-314.

Beucher, S. (2010). Image segmentation and Mathematical morphology. [Online] cited 5 March 2011, Available online as .pdf at: http://cmm.ensmp.fr/~beucher/wtshed.html#mark.

Blaschke, T. and J. Strobl, (2001). "What's wrong with pixels? Some recent developments interfacing remote sensing and GIS." *GIS – Zeitschrift für Geoinformationssysteme,* Vol. 6, No. 1, pp. 12-17.

Blaschke, T. (2010). "Object based image analysis for remote sensing." *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 65, No. 1 pp. 2-16.

Blaschke, T., C. Burnett, and A. Pekkarinen (2006). "Image segmentation methods for object based analysis and classification. In: *Remote Sensing and Digital Image Analysis. Including the Spatial Domain*. Book Series: Remote Sensing and Digital Image Processing, Eds. de Jong, S.M., and F. D. van der Meer, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 211-236.

Bouman, C.A. and M. Shapiro (1994). "A multiscale random field model for Bayesian image segmentation." *IEEE Transactions on Image Processing,* Vol. 3, No. 2, pp. 162-177.

Carleer, A.P., O. Debeir, and E. Wolff (2005). “Assessment of very high spatial resolution satellite image segmentations.” *Photogrammetric Engineering & Remote Sensing,* Vol. 71, No. 11, pp. 1285-1294.

Castilla, G., G. J. Hay, and J. R. Ruiz-Gallardo (2008). “Size-constrained Region Merging (SCRM): An Automated Delineation Tool for Assisted Photointerpretation.” *Photogrammetric Engineering & Remote Sensing*, Vol. 74, No. 4, pp. 409-419.

Chen, Z., Z. Zhao, P. Gong, and B. Zeng (2006). “A new process for the segmentation of high resolution remote sensing imagery.” *International Journal of Remote Sensing,* Vol. 27, No. 22, pp. 4991.

Cheng, H.D., X. H. Jiang, Y. Sun, and J. Wang (2001). “Color image segmentation: advances and prospects.” *Pattern Recognition,* Vol. 34, No. 12, pp. 2259-2281.

Corcoran, P., A. Winstanley, and P. Mooney, (2010). “Segmentation performance evaluation for object-based remotely sensed image analysis.” *International Journal of Remote Sensing*, Vol. 31, No. 3, pp. 617-645.

Costa, G.A.O.P., R. Q., Feitosa, T. B. Cazes, and B. Feijó, (2008). “Genetic adaptation of segmentation parameters.” In *Object-Based Image Analysis,* Blaschke T., S. Lang and G.J. Hay (Eds.), Springer-Verlag, Berlin: Hiedelberg, Germany, pp. 679-695

Definiens AG, 2009. eCognition Developer 8 User Guide, published by *Definiens AG, Germany.*

D'Elia, C., G. Poggi, and G. Scarpa (2003). “A tree-structured Markov random field model for Bayesian image segmentation.” *IEEE Transactions on Image Processing,* Vol. 12, No. 10, pp. 1259-1273.

Drăgut, L., D. Tiede, and S. R. Levick, (2010). “ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data.” *International Journal of Geographic Information Science,* Vol. 24, No. 6, pp. 859-871.

Estes, J.E., 1999, Lecture 2: Elements, Aids, Techniques and Methods of Photographic/Image Interpretation. [Online] cites 21 July, 2010, Available online at: http://userpages.umbc.edu/~tbenja1/umbc7/santabar/vol1/lec2/2lecture.html.

Falkowski, M. J., M. A. Wulder, J. C. White, and M. D. Gillis, (2009). “Supporting large-area, sample-based forest inventories with very high spatial resolution satellite imagery.” *Progress in Physical Geography,* Vol. 33, No. 3, pp. 403-423.

Fan, J., M. Han, and J. Wang (2009). “Single point iterative weighted fuzzy C-means clustering algorithm for remote sensing image segmentation.” *Pattern Recognition,* Vol. 42, No. 11, pp. 2527-2540.

Freixenet, J., X. Muñoz, D. Raba, J. Martí, and X. Cufí (2002). "Yet Another Survey on Image Segmentation: Region and Boundary Information Integration." In *Lecture Notes in Computer Science, Computer Vision — ECCV 2002,*Eds. Goos, G., J. Jartmanis, and J. van Leeuwen, Springer-Verlag, Berlin, Germany, Vol. 2352, pp. 408-422.

Fu, K.S. and J. K. Mui (1981). "A survey on image segmentation." *Pattern Recognition,* Vol. 13, No. 1, pp. 3-16.

Geman, S. and D. Geman (1984). "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images." *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 6, No. 6, pp. 721-741.

Gonzalez, R.C. and R. E. Woods (2002). *Digital Image Processing*, 2nd ed., Prentice-Hall, New Jersey, USA.

Guindon, B. (1997). "Computer-based aerial image understanding: a review and assessment of its application to planimetric information extraction from very high resolution satellite images. *Canadian Journal of Remote Sensing,* Vol. 23, No. 1, pp. 38-47.

Guo, D., Atluri, V. and Adam, N., (2005). "Texture-Based Remote-Sensing Image Segmentation." In *IEEE International Conference on Multimedia and Expo, 2005, (ICME '05)*, 6-8 July 2005, Amsterdam, The Netherlands, pp. 1472-1475.

Haralick, R.M. and L. G. Shapiro (1985). "Image segmentation techniques." *Computer Vision, Graphics, & Image Processing,* Vol. 29, No. 1, pp. 100-132.

Hasanzadeh, M. and S. Kasaei (2010). "A multispectral image segmentation method using size-weighted fuzzy clustering and membership connectedness." *IEEE Geoscience and Remote Sensing Letters,* 7, No. 3, pp. 520-524.

Hay, G. J. and G. Castilla (2006). "Object-Based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT)." In 1st *International Conference on Object-based Image Analysis* (OBIA 2006), IAPRS *XXXVI* – 4/C42, 4-5 July, Salzburg, Austria, pp. 1-3.

Hay, G.J., and D. J. Marceau (2004). "Multiscale object-specific analysis (MOSA): an integrative approach for multiscale landscape analysis." In: *Remote Sensing and Digital Image Analysis. Including the Spatial Domain*. Book Series: Remote Sensing and Digital Image Processing, Eds. de Jong, S.M., and F. D. van der Meer, Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 71–92

Hay, G.J., T. Blaschke, D. J. Marceau, and A. Bouchard (2003). "A comparison of three image-object methods for the multiscale analysis of landscape structure." *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 57, No. 5-6, pp. 327-345.

Herold, M., M. E. Gardner, and D. A. Roberts (2003). "Spectral resolution requirements for mapping urban areas." *IEEE transactions on Geoscience and Remote Sensing*, Vol. 41, No. 9, pp. 1907-1919.

Hester, D. B., S. A. C. Nelson, H. L. Cakir, S. Khorram, and H. Cheshire (2010). "High-resolution land cover change detection based on fuzzy uncertainty analysis and change reasoning." *International Journal of Remote Sensing,* Vol. 31, No.2, pp. 455-475

Hu, X., C. V. Tao, and B. Prenzel (2005). "Automatic segmentation of high-resolution satellite imagery by integrating texture, intensity, and color features." *Photogrammetric Engineering and Remote Sensing,* Vol. 71, No. 2, pp. 1399-1406.

ITT Visual Information Solutions, 2011. Customizing ENVI with IDL Whitepaper, published by *ITT Visual Information Solutions*.

Ju, J., S. Gopal, S. and E. D. Kolaczyk (2005). "On the choice of spatial and categorical scale in remote sensing land cover classification." *Remote Sensing of Environment,* Vol. 96, No. 1, pp. 62-77.

Karantzalos, K., and D. Argialas (2009). "A region-based level set segmentation for automatic detection of Man-made objects from aerial and satellite images." *Photogrammetric Engineering and Remote Sensing,* Vol. 75, No. 6, pp. 667-677.

Kim, J.-B., and H. –J. Kim (2003). "Multiresolution-based watersheds for efficient image segmentation." *Pattern Recognition Letters,* Vol. 24, No. 1-3, pp. 473-488.

Lang, S., E. Schöpfer, and T., Langanke (2009). "Combined object-based classification and manual interpretation-synergies for a quantitative assessment of parcels and biotopes." *Geocarto International*, Vol. 24, No.2, pp. 99-114.

Li, N., H. Huo, and T. Fang (2010). "A novel texture-preceded segmentation algorithm for high-resolution imagery." *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 48, No.7, pp. 2818-2828.

Li, P. and X. Xiao (2007). "Multispectral image segmentation by a multichannel watershed-based approach." *International Journal of Remote Sensing,* Vol. 28, No. 19, pp. 4429-4452.

Li, S. Z. (2009). *Markov Random Field Modeling in Image Analysis,* 3rd edition, Springer-Verlag, London, England, pp. 1-10: 199-210.

Lindberg, T. (1996). "Scale-space: A framework for handling image structures at multiple scales." In: Proc. CERN School of Computing, Egmond aan Zee, The Netherlands, 8–21 September, pp. 1-12.

Lizarazo, I., and J. Barros, 2010. Fuzzy Image Segmentation for Urban Land-Cover Classification, Photogrammetric Engineering & Remote Sensing, Vol. 76, No. 2, pp. 151-162.

Marpu, P.R., M. Neubert, H. Herold, and I. Niemeyer (2010). “Enhanced evaluation of image segmentation results.” *Journal of Spatial Science,* Vol. 55, No. 1, pp. 55-68.

Mas, J.F. and J.J. Flores (2008). “The application of artificial neural networks to the analysis of remotely sensed data.” *International Journal of Remote Sensing,* Vol. 29, No. 3, pp. 617-663.

Mather, P.M., (2004). “Pattern recognition using artificial neural networks.” In *Computer Processing of Remotely-Sensed Images: An Introduction,* 3rd Ed., John-Wiley and Sons, Chichester, England, pp. 224-228.

Maxwell, T. L. and Y. Zhang (2005). “A fuzzy approach to supervised segmentation parameter selection for object-based classification.” In Proceedings of SPIE: Applications of Digital Image Processing XXVIII, 2-4 Aug., San Diego, USA, Vol. 5909, pp. 528-538. Mayer, H., S. Hinz, U. Bacher, E. Baltsavias (2006). “A test of automatic road extraction approaches.” In*: Photogrammetry Computer Vision and Image Analysis, 20-22 Sept. Bonn, Germany, IAPRS’06*, Vol. XXXVI-3, pp. 209–214.

McGuiness, K., and N. E. O’Connor (2010). “A comparative evaluation of interactive segmentation algorithms.” *Pattern Recognition*, Vol. 43, No.2, pp. 434-444.

Mesev, V. (2003). “Remotely sensed cities: an introduction.” In *Remotely Sensed Cities,* Taylor and Francis, London, UK, pp. 1-19.

Meyer, F. and S. Beucher (1990). “Morphological segmentation.” *Journal of Visual Communication and Image Representation*, Vol. 1, No.1, pp. 21-46.

Mishra, R. K. And Zhang, Y. (2011). “ Object-oriented approach for remote sensing image processing using Java advanced imaging.” Proceedings of the *32nd Canadian Symposium on Remote Sensing,* Sherbrooke, Québec, Canada, 13-16 July 2011.

Mohammadzadeh, A., and A. J. V. Zoej (2010). “A Self-organizing Fuzzy Segmentation (SOFS)Method for Road Detection from High Resolution Satellite Images.” *Photogrammetric Engineering & Remote Sensing*, Vol. 76, No.1, pp. 27-35.

Neubert, M., H. Herold and G. Meinel, G. (2008). “Assessing image segmentation quality-concepts, methods and applications.” In: *Object Based Image Analysis*, Eds. Blaschke T., S. Lang and G.J. Hay, Springer-Verlag, Berlin, Germany, pp. 760–784.

Ojala, T. and M. Pietikäinen (1999). “Unsupervised texture segmentation using feature distributions.” *Pattern Recognition,* Vol. 32, No.3, pp. 477-486.

Pal, N.R. and S. K. Pal (1993). “A review on image segmentation techniques.” *Pattern Recognition,* Vol. 26, No. 9, pp. 1277-1294.

Peng, J., D. Hang, and Y. Liu, (2005). “An improved snake model for building detection from urban aerial images.” *Pattern Recognition Letters*, Vol. 26, No.5, pp. 587-595.

Pesaresi, M. and J.A. Benediktsson (2001). “A new approach for the morphological segmentation of high-resolution satellite imagery.” *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 39, No.2, pp. 309-320.

Pham, D. L., C. Xu, J. L. Prince (2000). “A Survey of Current methods in medical image segmentation.” *Annual Review of Biomedical Engineering*, Vol. 2, No.1, pp. 315-337.

Platt, R. V., and L. Rapoza (2008). “An evaluation of an object-oriented paradigm for land use/land cover classification.” *The Professional Geographer*, Vol. 60, No.1, pp. 87-100.

Poggi, G., G. Scarpa, and J. B. Zerubia (2005). “Supervised segmentation of remote sensing images based on a tree-structured MRF model.” *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 43, No.8, pp. 1901-1911.

Reed, R.T. and J.M.H. Du Buf (1993). “A review of recent texture segmentation and feature extraction techniques.” *Computer Vision Graphics and Image Processing: Image Understanding,* Vol. 57, No.3, pp. 359-372.

Schiewe, J. (2002). “Segmentation of high-resolution remotely sensed data-concepts, applications and problems.” In *ISPRS Technical Commission IV symposium: Geospatial theory, processing and application XXXIV part 4,* 9-12 July 2002, Ottawa, Canada, pp. 358-363.

Shackelford, A.K. and C.H. Davis (2003). “A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas.” *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 9, pp. 1920-1932.

Shankar, B. U. (2007). “Novel classification and segmentation techniques with application to remotely sensed images.” In Vol. 4400, *Lecture Notes on Computer Science: Transactions on Rough Set VII*, Ed. J. F. Peters, A. Skowron, V. W. Marek, E. Orlowska, R. Slowinski, and W. Ziarko. Springer-Verlag, Berlin, Germany, pp. 295-380.

Spirkovska, L. (1993). A summary of image segmentation techniques. [Online] cited 10 Sept. 2010. Available online as .pdf at: http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19940006802_1994006802.pdf.

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,* Vol. 28, No. 20, pp. 4625-4644.

Tizhoosh, H.R. and H. Haussecker (2000). "Fuzzy image processing: An overview." In *Computer Vision And Applications: A Guide for Students and Practitioners,* Eds. Jähne B. and H. Haussecker, Academic Press, Boston, USA, pp. 541-576.

Triaz-Sanz, R., G. Stamon and J. Louchet (2008). "Using color, texture, and hierarchical segmentation for high-resolution remote sensing." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 63, No.2, pp. 156-168.

Tso, B., and P.M. Mather (2004). "Pattern recognition using artificial neural networks." In *Classification methods for remotely sensed data,* Taylor & Francis, London, UK,. pp. 102-140

Wang, D., (1997). "A multiscale gradient algorithm for image segmentation using watershelds." *Pattern Recognition,* Vol. 30, No. 12, pp. 2043-2052.

Wang, Z., J. R. Jensen, and J. Im (2010). "An automatic region-based image segmentation algorithm for remote sensing applications." *Environmental Modelling and Software,* Vol. 25, No. 10, pp. 1149-1165.

Werff van der, H. M. A. and F. D. Meer van der (2008). "Shape-based classification of spectrally identical objects." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 63, No.2, pp 251-258.

Woodcock, C.E., and A. H. Strahler (1987). "The factor of scale in remote sensing." *Remote Sensing of Environment,* Vol. 21, No. 3, pp. 311-332.

Wuest, B., and Y. Zhang (2008). "Supervised Region Based Segmentation of QuickBird Multispectral Imagery." In *Proceedings of the 2008 IEEE International Geoscience & Remote Sensing Symposium*, IGARSS 2008, 7-11 July, Boston, USA, pp. 1021-1024.

Wuest, B., and Y. Zhang (2009). "Region based segmentation of QuickBird multispectral imagery through band ratios and fuzzy comparison." *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 64, No. 1, pp. 55-64.

Xia, Y. and D. Feng (2009). "A General Image Segmentation Model and its Application." In *Fifth International Conference on Image and Graphics,* ICIG '09, 20-23 Sepember 2009*,* Shanxi, China, pp. 227-231.

Zadeh, L.A. (1965). "Fuzzy Sets." *Information and Control,* Vol. 8, No. 3, pp. 338-353.

Zadeh, L.A. (1975). "The Concept of a Linguistic Variable and its Application to Approximate Reasoning-I." *Information Sciences,* Vol. 8, No. 3, pp. 199-249.

Zhang, H., J. E. Fritts, and S. A. Goldman (2008). “Image segmentation evaluation: A survey of unsupervised methods.” *Computer Vision and Image Understanding,* Vol. 110, No. 2, pp. 260-280.

Zhang, Y.J. (1997). “Evaluation and comparison of different segmentation algorithms.” *Pattern Recognition Letters,* Vol. 18, No. 10, pp. 963-974.

Wu, Z., L. Yin. L., and M. Qin (2009). “Granular Approach to Object-Oriented Remote Sensing Image Classification.” In *Lecture Notes in Artificial Intelligence,* Vol. 5589*,* Ed., R. Goebel, Springer-Verlag, Berlin/Hiedelberg, Germany, pp. 563-570.

Zhou, Y. and Y. Q. Wang (2008). “Extraction of impervious surface areas from high spatial resolution imagery by multiple agent segmentation and classification.” *Photogrammetric Engineering and Remote Sensing*,Vol. 74, No. 7, pp. 857-868.

Zouagui, T., H. Benoit-Cattin, and C. Odet (2004). “Image segmentation functional model.” *Pattern Recognition,* Vol. 37, No. 9, pp. 1785-1795.

CHAPTER 3

A SUPERVISED METHODOLOGY FOR OPTIMAL PARAMETER ESTIMATION OF MULTIRESOLUTION SEGMENTATION WITHIN ECOGNITION

This chapter contains a journal paper to be submitted to an international journal and referred as:

Dey, V., Y. Zhang, and M. Zhong (2011). "A Supervised Methodology for Optimal Parameter Estimation of Multiresolution Segmentation within eCognition." *International Journal of Remote Sensing,* (to be submitted).

The presented paper identifies the problems of multiresolution segmentation (one of the most widely-used VHR image segmentation technique, as established in Chapter 2) and attempts to solve those problems. In addition, it also justifies the proposed solution with experiments on different VHR images. In order to present clearly, the original paper has been slightly edited.

Abstract

Object-based image analysis (OBIA) has emerged as one of the major drivers of remote sensing applications in the last decade. Image Segmentation is regarded as both the fundamental and most critical step of OBIA. Fortunately, several software packages,

both commercial and open source, exist to solve the problem of image segmentation. Among existing software, eCognition™ has emerged as one of the major leader in OBIA, especially because of its multiresolution segmentation and hierarchical classification functionality. However, multiresolution segmentation faces the problem of parameter optimization. The traditional solution of parameter optimization is based on trial and error approach, which is time consuming. To solve this problem, this paper introduces a supervised methodology to estimate the optimal parameters of the multiresolution segmentation in eCognition™. The optimization is devised using a heuristics approach, where simple object features such as, standard deviation, mean difference to neighbours, and compactness are used. The whole implementation is within framework of eCognition™ using customized object features, and hierarchical processing rules. In addition, the approach proposes a global segmentation evaluation method as well as a guideline to select an appropriate training object, which is required for supervised solution. The experiments demonstrate that the proposed approach based parameters are as effective as obtained by trial and error approach. However, the proposed approach estimates the effective parameter within five minutes for any land cover, which is generally not possible with trial and error based approach. Hence, the major contribution of the proposed approach is the reduction of time in estimation of effective parameters for multiresolution segmentation.

3.1 Introduction

The development of object-based image analysis (OBIA) has been widely recognized in remote sensing (RS) [Blaschke, 2010]. This recognition is primarily attributed to the growth of very high spatial resolution (VHR) satellite images (Spatial resolution <= 1m) in the last decade [Blaschke and Strobl, 2001; Blaschke, 2010]. However, OBIA needs a meaningful identification of the land cover objects of an image. A land cover object is a homogeneous group of pixels (also known as a region) with a meaning in the real world, where meaning is defined by the geographic classes of an image [Hay et al., 2003; Blaschke et al., 2006]. For example, regions which are enclosed by the red-coloured polygons in Figure 3.1 correspond to a real world land cover class of buildings. These regions (also known as segments) are identified by image segmentation process, which is the fundamental step of OBIA [Burnett and Blaschke, 2003; Benz et al., 2004; Lang et al., 2009; Blaschke, 2010]. Therefore, image segmentation has a major impact on the quality of the overall results of OBIA applications, such as land cover classification, feature extraction, and change detection.

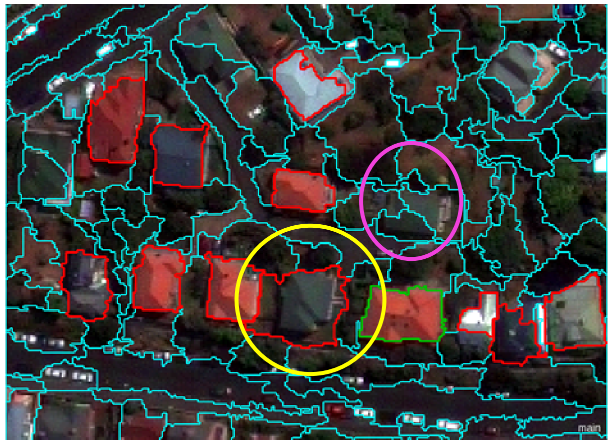


Figure 3.1: Demonstrates the meaning of an object, where the polygons (in red) are the objects/segments representing buildings of a real world. The objects which are encircled in pink show a non-meaningful representation of a building. The objects which are encircled in pink and in yellow demonstrate over-segmentation and under-segmentation respectively.

3.1.1 Popularity of Multiresolution Segmentation

In RS, multi-scale analysis based image segmentation techniques are among the best techniques because they aim to analyze different land cover objects at their best scales of analysis [Hay et al., 2003; Benz et al., 2004; Blaschke, 2010]. A multi-scale segmentation technique based OBIA has been first implemented in the commercial software eCognition™ (now known as eCognition™ Developer and owned by Trimble Inc.) in the year 2000 [Baatz and Schäpe, 2000; Flanders et al., 2003]. This multi-scale technique is known as multiresolution segmentation in eCognition™ and is based on fractal net evaluation approach (FNEA) [Baatz and Schäpe, 2000].

In addition, eCognition^{TM} is the most widely used software for OBIA applications. This is based on the comprehensive review of Blaschke [2010], which comprised of review of approximately 800 research papers on segmentation. He found that eCognition^{TM} based segmentation was utilized in more than 50-55% of the reviewed papers and the rest used several individual segmentation algorithms, such as Wuest and Zhang [2009] and Lizarazo and Barros [2010]. Moreover, on a comparison of more than ten different software, Neubert et al. [2008] and Marpu et al. [2010] concluded that eCognition^{TM} is among the best available segmentation software in general.

The success of eCognition^{TM} led to the development of few more software with OBIA capability, such as Feature-Analyst^{TM} in 2001 (now owned by Overwatch Textrons), ENVI Fx in 2007 (by ITT Visual Information Solutions), FeatureObjeX in 2008 (by PCI-Geomatics^{TM}), and IMAGINE Objective in 2008 (by ERDAS^{TM}). However, eCognition^{TM} is still the front runner with respect to OBIA based applications [Blaschke 2010].

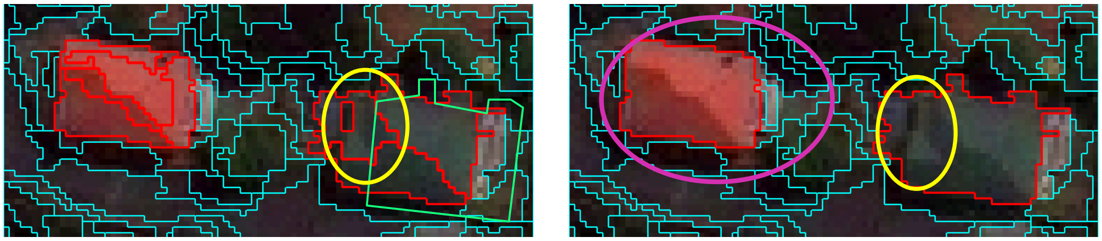
3.1.2 Problems of Multiresolution Segmentation

Although eCognition^{TM}'s multiresolution segmentation is popular, the segmentation results are sensitive with respect to its three parameters: Scale value, Shape weight, and Compactness weight [Hay et al., 2003; Maxwell, 2005; Tian and Chen, 2007; Marpu et al., 2010]. Among the three, the most sensitive parameter is Scale value [Definiens AG, 2009]. The common strategy of estimation of the parameters is the selection of a set of parameters and test them using trial and error process, until the results are either visually pleasing to the operator or the operator discontinues the process [Flanders et al., 2003; Maxwell, 2005; Platt and Rapoza, 2008]. Essentially, this is not a conceptually strong process to estimate the effective parameters with this strategy.

The trial and error strategy is also time consuming and fails to guarantee a consistent solution to the problem of parameter estimation. Moreover, for a faster solution, this strategy requires operator's experience with the image and the segmentation algorithm. Therefore, the major problem related to parameter estimation is the uncertain/long duration of the trial and error strategy [Hay et al., 2003; Marpu et al., 2010].

3.1.3 Review of Existing Solution

Several researchers have proposed solutions to overcome the problem of parameter estimation of multiresolution segmentation. For example, Maxwell [2005] proposed a fuzzy based supervised approach to estimate the three parameters. The approach used a multi-level segmentation where a training object is defined by merging the sub-objects of a land cover class as shown in Figure 3.2. The approach mapped spectral, texture, shape, and size based properties of the training object and its sub-objects using fuzzy inference systems (FISs) to obtain effective parameters. The approach estimates the effective solution faster than the trial and error approach as demonstrated by Zhang et al., [2010].



[a] Initial over segmentation of the possible training objects shown in red, where yellow circle shows the sub-objects which include significant number of pixels outside the desired boundary (green) of the training object

[b] Correct and incorrect training object (encircled in pink and yellow respectively) selection from the merging of the sub-objects of the initial over-segmented image.

Figure 3.2: Illustrates the correct and incorrect training objects [b] and its sub-objects

On the other hand, Costa et al. [2008] used a genetic algorithm to estimate the parameters. The fitness function of the genetic algorithm utilized a segmentation quality evaluation measure to estimate the effective parameters. The measure compared the geometry of the selected reference objects and the segments obtained for the reference objects to determine the effectiveness.

Tian and Chen [2007] also used a segmentation evaluation measure. However, instead of using any algorithm to estimate a set of three parameters, they defined a sequence of the sets. These sets were used to obtain different segmentation results. These results were compared based on a segmentation evaluation measure and the parameter set with the most effective results is identified as the effective parameters. Marpu et al. [2010] used a strategy similar to Tian and Chen [2007]. However, the segmentation quality evaluation measure and the sequence of parameter sets were different from Tian and Chen [2007].

The above mentioned three approaches require a long time (in hours) to estimate the effective parameters. Although these approaches have certainty of an effective estimation, the major problem related to the duration of estimation of effective parameters still remains.

Recall from the Section 3.1.2, the scale parameter is the most sensitive parameter. Considering this fact, Huang et al. [2003], Möller et al. [2007], Kim et al. [2008], Chen et al. [2009], and Drăgut et al. [2010] suggested different approaches to estimate only the effective scale parameter of different land cover objects in a VHR image. However, for an effective multiresolution segmentation, a proper shape weight parameter is also essential [Tian and Chen, 2007].

3.1.4 Methodology of the Proposed Approach

Similar to Maxwell [2005], the proposed supervised approach uses a training process for the parameter estimation. The training process uses an object of interest of a particular land cover class (hereby called as the training object), which is identified by merging its sub-objects. These sub-objects are obtained at an appropriate lower level over-segmented image as demonstrated in Figure 3.2. The idea of merging the sub-objects to form the training object is in accordance with a hierarchical representation of objects within eCognition™ [Benz et al., 2004]. However, the proposed approach uses a crisp logic instead of fuzzy logic used by Maxwell [2005].

The approach also utilizes five features of the training objects and its sub-objects to estimate the effective parameters. These features are brightness, mean difference to neighbours, compactness, size, and texture. The estimated parameters are used to segment the entire image to identify all the objects of interest of the selected land cover class. The segmentation process using the proposed approach is demonstrated with Figure 3.1. In the Figure, the training object is a building land cover object (in green) and the objects of interest (in red) class are the objects of building class. These objects are identified after the segmentation with the estimated parameters.

The above-mentioned procedure of the approach is proposed for a single land cover class extraction. However, the process can easily be extended for hierarchical land cover classes of an image by using different training objects of the parent and child class at different scales. For example, an effective segmentation of a Tree cover class would serve as the over-segmented image for the Forest cover class, which is the parent class of the Tree cover class. Hence, different training objects for the two hierarchical classes can be selected at two different levels of segmentation.

This paper also proposes a global segmentation results evaluation technique. The technique is based on the detection of under-segmentation and over-segmentation of the selected reference segments. This evaluation technique can also be used to obtain an appropriate training object of a land cover class. Further, the technique is used to establish the effectiveness of solution with respect to the parameters estimated from trial and error based approach.

3.1.5 The Contribution of this Paper

This paper proposes a supervised approach for finding the three effective parameters of multiresolution segmentation. The approach specifically addresses the major problem of trial and error based approach as defined in the section 3.1.2. The duration problem of multiresolution segmentation (see Section 3.1.2) is reduced by implementing the approach using the features and functions of eCognition™ Developer 8.0. In any case of the parameter estimation, the proposed approach should identify the effective parameters within five minutes. Moreover, the effectiveness of the proposed approach is established using the proposed evaluation technique. Hence, this paper proposed a parameter estimation approach which is fast, easy to implement, and effective.

To provide with the background information, this paper will first briefly explain the concept of multiresolution segmentation in the Section 3.2. The proposed supervised approach and its implementation steps will be detailed in the Section 3.3. The Section 3.4 deals with the comparison of segmentation results of VHR images of different sensors and locations. These results are obtained from the estimated parameters. The comparison analysis Section will be followed by the conclusions in the Section 3.5.

3.2 Multiresolution Segmentation Algorithm

The multiresolution segmentation algorithm is a widely used algorithm employed within the commercial software eCognition^{TM} [Baatz and Schäpe, 2000]. The algorithm is a bottom-up approach (since it starts with a single seed pixel) and follows a pair-wise region merging process. The aim of this algorithm is to minimize the heterogeneity of the image objects obtained after the pair-wise region merging [Benz et al., 2004].

In the segmentation process, at each step, the heterogeneity of a pair of adjacent image objects is evaluated. The pair is merged based on the two conditions: (1) if the merging is local mutual best fitting, i.e., the heterogeneity of the merging of the selected pair is minimum out of all possible merging pairs associated with any one of the two objects; and (2) if the heterogeneity due to the merging of the selected pair is less than the square of a scale parameter threshold (S) [Baatz and Schäpe, 2000]. The second condition justifies the analysis of the Section 3.1.2 related to the scale parameter's highest sensitivity towards segmentation results. The next Section provides details of the heterogeneity equations used for the pair-wise region merging of the multiresolution algorithm.

3.2.1 The Measures of Heterogeneity Change

The heterogeneity of multiresolution segmentation algorithm is composed of spectral and shape heterogeneity, $h_{spectral}$ and h_{shape}. These heterogeneities are the differences in the size-weighted spectral and shape features of the two adjacent objects and their merged object. The overall spectral heterogeneity is the sum of the spectral heterogeneity for each layer and the overall heterogeneity is defined as:

$$h_{spectral} = \sum_{i}^{m} w_i \left(n_{Obj1+Obj2} \cdot \sigma_i^{Obj1+Obj2} - \left(n_{Obj1} \cdot \sigma_i^{Obj1} + n_{Obj2} \cdot \sigma_i^{Obj2} \right) \right) \qquad (3.1)$$

where m represents the number of layers of the image chosen for segmentation, w_is are the user-defined weights associated with the layer i, $Obj1$ and $Obj2$ represent the two adjacent objects selected for merging, n is the number of pixels or size of the objects, $Obj1+ Obj2$ represents the merged object resulting from the merging of the two adjacent objects $Obj1$ and $Obj2$, and σ_i is the standard deviation of the objects of the layer i [Benz et al., 2004].

While the spectral heterogeneity uses only a single spectral feature (standard deviation), the overall shape heterogeneity, h_{shape}, is composed of the weighted average of the two shape features heterogeneity: compactness feature heterogeneity, $h_{compactness}$, and smoothness feature heterogeneity, $h_{smoothness}$. h_{shape} is defined as:

$$h_{shape} = w_{compactness} \cdot h_{compactness} + (1 - w_{compactness}) \cdot h_{smoothness} \tag{3.2}$$

where $w_{compactness}$ is the user-defined weight of the compactness feature heterogeneity change (range 0 to 1) [Benz et al., 2004]. The compactness feature of an object is defined as:

$$Compactness = \frac{l_{Obj}}{\sqrt{n_{Obj}}} \tag{3.3}$$

where l_{obj} is the perimeter of the object [Benz et al., 2004].

On the other hand, the smoothness feature of an object is defined as:

$$Smoothness = \frac{l_{Obj}}{b_{obj}} \tag{3.4}$$

where b_{obj} is the perimeter of the object's bounding box [Benz et al., 2004]. The compactness shape feature represents how the pixels are distributed around the centroid of the object, e.g., circle is the most compact object. The smoothness shape feature refers to the fluctuations or smoothness of the border of the object, e.g., rectangle is smoothest object for a raster image. In the raster analysis, the most compact object is represented by a square (different from circle of the vector analysis) but the smoothest object remains the same. Based on the compactness feature, the compactness heterogeneity change, $h_{compactness}$, is defined by:

$$h_{compactness} = n_{Obj1+Obj2} \cdot \frac{l_{Obj1+Obj2}}{\sqrt{n_{Obj1+Obj2}}} - \left(n_{Obj1} \cdot \frac{l_{Obj1}}{\sqrt{n_{Obj1}}} + n_{Obj2} \cdot \frac{l_{Obj2}}{\sqrt{n_{Obj2}}} \right) \quad (3.5)$$

where n_{Obj} is the number of pixels/size of the objects [Benz et al., 2004].

Similarly, the smoothness heterogeneity, $h_{smoothness}$, is defined as:

$$h_{smoothness} = n_{Obj1+Obj2} \cdot \frac{l_{Obj1+Obj2}}{b_{Obj1+Obj2}} - \left(n_{Obj1} \cdot \frac{l_{Obj1}}{b_{Obj1}} + n_{Obj2} \cdot \frac{l_{Obj2}}{b_{Obj2}} \right) \quad (3.6)$$

where b_{obj} is the perimeter of the object's bounding box [Benz et al., 2004]. The overall heterogeneity change due to the potential merging of the two adjacent objects is the weighted sum of $h_{spectral}$ and h_{shape} and is defined as 'merging cost' (merging threshold), M_c, by Baatz and Schäpe [2000]. The overall heterogeneity change is formulated as:

$$M_c = (1 - w_{shape}) \cdot h_{spectral} + (w_{shape}) \cdot h_{shape} \quad (3.7)$$

where w_{shape} is the user-defined weight parameter assigned to the shape heterogeneity change [Benz et al., 2004]. The range of w_{shape} lies in between 0.0 to 0.9. The limit of M_c is determined by the scale parameter threshold, S.

The two adjacent objects in the pair-wise region merging process qualifies for the final merging if $M_c < S^2$ and the criterion of local mutual best fitting is satisfied [Baatz and Schäpe, 2000]. The pair-wise merging of the segmentation follows an iterative process, where the iteration stops if the heterogeneity of the merging of every possible two adjacent object pair in the entire image exceeds the scale threshold. During each of the iterations of the segmentation process, a distributed treatment order is followed for the final merging candidates: (1) to achieve uniform growth of objects with similar scales of heterogeneity; and (2) to ensure the repeatability of segmentation with the same parameters. The repeatability of multiresolution is its special feature because this feature is lacked by many other segmentation algorithms [Tian and Chen, 2007].

The merging process also requires four user-defined parameters as defined by the Equations (3.1), (3.2), (3.5), (3.6), and (3.7). These user-defined parameters are: *S*, w_{shape}, $w_{compactness}$, and w_i (i = 1 to n). However, the major problem lies with the estimation of first three parameters and the parameter w_is are assumed to be 1 successfully [Hofmann, 2001].

3.3 Methodology

As described in the last Section, the result of multiresolution segmentation algorithm depends on the three parameters: Scale value, Shape weight, and Compactness weight. The scale parameter is the most important parameter because it governs the average size of the segments. Consequently, the effective scale value for the same object in images of different spatial resolutions increases with the resolution of the image [Castilla et al., 2008]. However, the scale is a unit less parameter. Hence, there is no direct mathematical relationship between the scale value and the size of the object in an image [Hay et al., 2003]. Therefore, the scale parameter is the most difficult parameter to estimate in multiresolution segmentation.

Although the major criteria of segmentation should rely on the spectral information, a suitable weight on the shape information improves the form of the shapes of the segments [Benz et al., 2004, Tian and Chen, 2007]. However, it is challenging to accurately estimate the shape weight parameter and the kind of shape information (compactness or smoothness) to be emphasized for effective segmentation [Tian and Chen, 2007].

As mentioned in the last two paragraphs, there are no mathematical relationships which can estimate the three parameters for effective segmentation. This is because the definition of effective segmentation lacks a general mathematical representation [Pal and Pal, 1993]. This suggests that the estimation of the parameters should rely on heuristic approaches, which are based on the concept of effective segmentation. The Section 3.1.3 has provided a brief introduction of such existing heuristic approaches.

The present paper also formulates a supervised heuristic approach whose property is defined in the above paragraph. At first, the paper identifies the criteria of effective segmentation. Then, the approach is formulated based on these criteria. The current Section describes the steps of the formulation of this approach.

3.3.1 Criteria for effective Segmentation

The aim of segmentation is to identify meaningful objects, where a meaningful object in a RS image should: (1) resemble to a land cover class in an appropriate shape and size; and (2) have the shape which visually satisfies a human operator [Tian and Chen, 2007; Castilla and Hay, 2008]. The definition of meaningful object concentrates on the local effectiveness of segmentation results, i.e., effective results for a selected object. On the other hand, globally effective segmentation results should neither have over-segmentation and nor under-segmentation. Therefore, effective segmentation results should identify meaningful objects such that the overall results are neither over-segmented nor under-segmented.

Over-segmentation in the present context is defined as the results where a significant number of the meaningful objects are over-segmented [Kim et al., 2008; Marpu et al., 2010]. Similarly, under-segmentation is defined as the results where a significant number of the meaningful objects are under-segmented [Kim et al., 2008; Marpu et al., 2010]. However, the over-segmentation results are preferred over under-segmentation results due to the possibility of handling the over-segmentation results during a segmentation-based classification process [Castilla and Hay, 2008].

As mentioned before, it is essential to visualize effective segmentation in terms of the meaningful objects delineation of a land cover class [Benz et al., 2004]. Figure 3.2

demonstrates both the over-segmented (in pink) and under-segmented (in yellow) results. The Figure also demonstrates an example of a non-meaningful shape and size delineation of a building object. This interpretation of effective segmentation is used to design the parameter estimation approach of this paper.

3.3.2 Estimation of the Parameters of Multiresolution Segmentation

Since effective segmentation requires information of land cover classes, the parameter estimation approach should incorporate this information. For example, Tian and Chen [2007] defined different criteria for the identification of meaningful objects of Roads and the Buildings land cover classes. Moreover, the parameter estimation approach should be capable of estimating different scale values for the identification of objects of different sizes [Flanders et al., 2003; Shackelford and Davis, 2003]. In addition, the identified objects should correspond to the reference objects as good as with traditional trial and error based approach. Overall, the proposed parameter estimation approach should be capable of estimating parameters: 1) for different land cover classes in a multi-level segmentation; and 2) for producing efficient segmentation based on the criteria of segmentation assessments.

The proposed approach of parameter estimation is designed to incorporate these requirements. The first requirement is satisfied by using a supervised approach, where different parameters are estimated based on training objects of different land cover classes. The second requirement is satisfied by using a global segmentation evaluation technique, which is proposed in this paper. Figure 3.3 depicts the general workflow of the proposed approach. The general terms which are used in the workflow are as described below:

1. Initial over-segmentation: A lower scale level segmentation result where all meaningful objects are over-segmented. For example, Figure 3.2[a] shows over-segmented results.
2. Training object: A segment formed by merging its sub-segments and corresponds to a land cover class. For example, Figure 3.2[b] shows a training object of a building land cover class.
3. Sub-objects: Small segments of the training object identified based on manual interpretation of the over-segmented results.
4. Customized object features: The feature values of an object which are formulated using the functions of eCognition™. The proposed approach uses three customized features (*Texture*, *Spectral stability,* and *Compactness*) that are defined by Maxwell [2005].

Essentially, the proposed approach aims to estimate the parameters such that the sub-objects of an initial over-segmented image merge to delineate the training object at higher level of segmentation. Due to this description of the aim, the training object can also be referred as the target object. The general steps of the workflow for a land cover class are described in the next few sub-Sections.

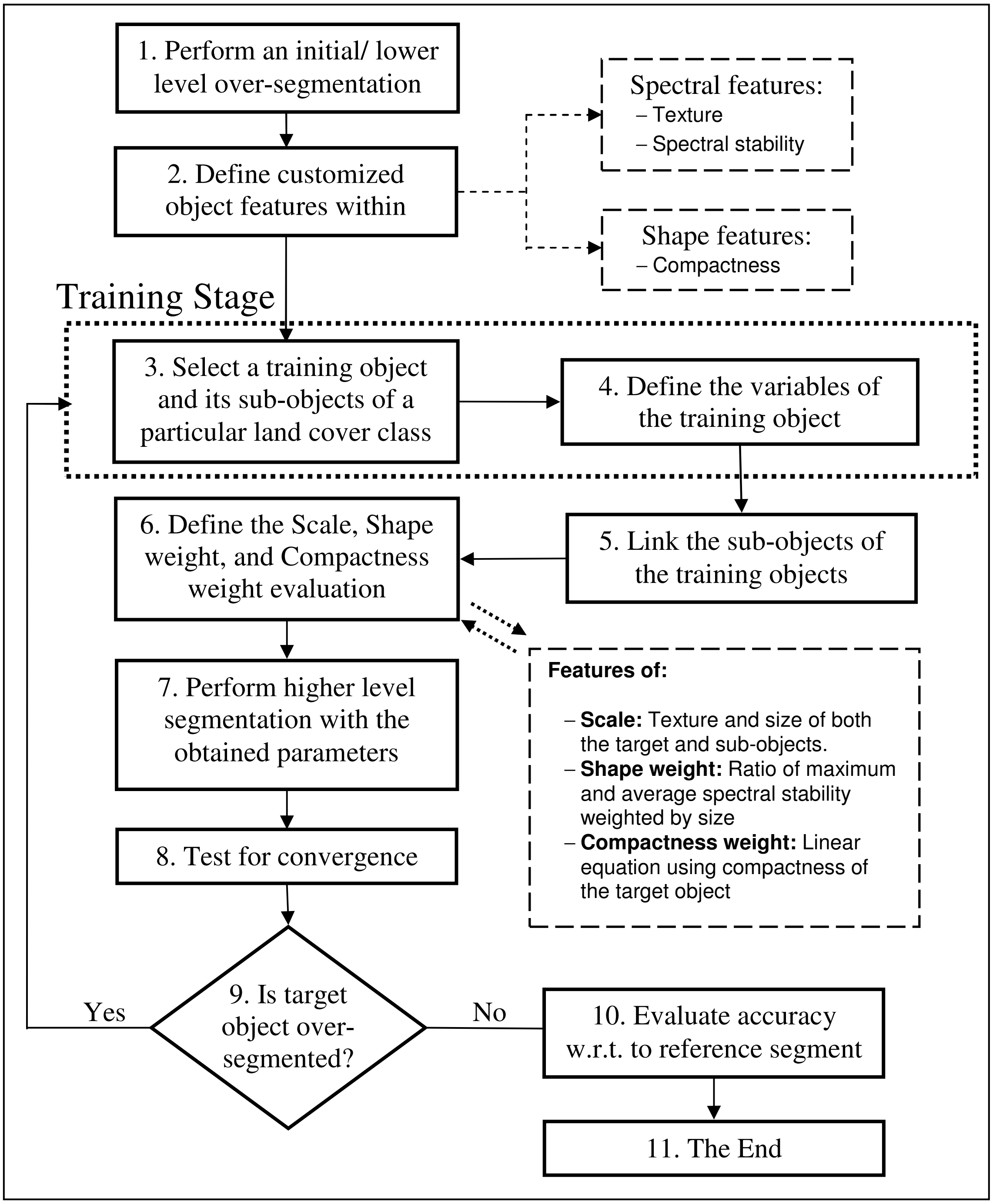


Figure 3.3: Workflow of the proposed parameter estimation approach. The values of the current Segmentation Parameter (Scale (S)), sub-objects information (*Texture, Spectral stability, Brightness, and Size*) and Target Object information (*Texture, Spectral stability, Brightness, Size, and Compactness*) are used in the approach to map target information with its sub-objects information in order to estimate the effective segmentation parameters ($w_{compactness}$, w_{shape}, and S) in an iterative manner.

3.3.2.1 Perform an Initial Over-segmentation using eCognition^{TM} (Step 1)

The initial over-segmentation of the training object should be conducted with: (1) a small scale parameter (*S*), e.g., default scale parameter value of multiresolution segmentation in eCognition^{TM}; (2) a little or no weight to shape parameter; and (3) equal weight to compactness and smoothness parameter. The parameters are intended to produce spectrally homogeneous objects.

3.3.2.2 Define Customized Object Features within eCognition^{TM} (Step 2)

The customized object features are defined by the operations on built-in spectral, shape, and size feature values for objects in eCognition^{TM}. These features and their equations are defined by Maxwell [2005] and are described in the next two sub-Sections

3.3.2.2.1 Spectral based customized object features

In multiresolution segmentation process, the spectral heterogeneity of an object is defined using the objects' standard deviation and size features (see Equation 3.1). Standard deviation of an object shows internal spectral variance of the object and is also a texture property of an object [Reed and Buf, 1993]. Therefore, Maxwell [2005] defined texture feature of an object as:

$$Texture(An\ object) = \frac{1}{m}\left[\sum_{i=1}^{m} \sigma_i^{obj}\right] \tag{3.8}$$

where m represents the number of spectral layers and σ_i^{obj} is the standard deviation of the spectral values of layer *i* of the object. Essentially, *Texture* represents internal heterogeneity of an object.

On the other hand, an external heterogeneity feature is defined by *Spectral stability.* This external heterogeneity feature represents the heterogeneity that will result from the merging of two adjacent objects. *Spectral stability* uses built-in *Mean_Difference_to_Neighbours* (MDN) feature of eCognition™. The MDN feature Δs_i^{obj} is defined as:

$$\Delta s_i^{obj} = \frac{1}{l} \cdot \sum_p [\, l_s^{obj_p} . \left| \bar{s}_i^{obj} - \bar{s}_i^{obj_p} \right|] \tag{3.9}$$

where *l* is the border length of the object; *p* represents the number of objects that are direct neighbours to the object; $l_s^{obj_p}$ is the length of shared border between the object and a direct neighbour object *p*; $\bar{s}_i^{obj}$ is the spectral mean value of a layer *i* for the object; and $\bar{s}_i^{obj_p}$ is the spectral mean value of layer *i* of the direct neighbour object *p* [Definiens AG, 2009].

Using the definition of MDN feature, the *Spectral stability* of an object is formulated by Maxwell [2005] as:

$$Spectral\ stability\ (An\ object) = \frac{1}{m}\sum_i^m \Delta s_i^{obj} \tag{3.10}$$

Essentially, *Spectral stability* refers to how spectrally stable the object is with respect to merging with its adjacent objects. In simple words, a low stability value favors the merging of the object with its adjacent objects and vice-versa.

3.3.2.2.2 Shape based customized object feature

The proposed approach uses only *Compactness* feature as shape based customized object feature. The choice of the *Compactness* feature is based on its use in the definition of shape heterogeneity (see Equation 3.2 and 3.5). In addition, the *Compactness* feature value is always 4x greater than the *Smoothness* feature value (see Appendix I for the proof). Hence, the *Smoothness* feature can be ignored here. The definition of *Compactness* feature is same as in Equation 3.3 [Definiens AG, 2009]. However, the built-in *Compactness* feature of eCognition™ is different from the definition of the *Compactness* feature in Equation 3.3 [Definiens AG, 2009]. This is why the feature is listed in the group of customized features.

Overall, the proposed approach for the parameter estimation needs the three customized features: *Texture* (internal spectral heterogeneity), Spectral *stability* (external spectral heterogeneity), and *Compactness* (shape heterogeneity). Apart from the customized features, the built-in features *Brightness* (average of spectral means of all spectral layers of an object) and *Size* (number of pixels of an object) are also employed for the parameter estimation.

3.3.2.3 Select a training object and its sub-objects (Step 3)

The proposed approach needs a training object of a land cover class to estimate the parameters. The training object is delineated by merging its sub-objects at a lower level of segmentation. Figure 3.2[a] and 3.2[b] illustrate how to select appropriate sub-objects for the selection of the training object of Building class.

3.3.2.4 Define the Variables of the Training Object (Step 4)

The variables of the training object represent the features of the training object. These variables remain fixed throughout a single iteration of the workflow. The names of these variables are: a) *TextureTO*, b) *CompactnessTO*, and c) *SizeTO*. Here, the suffix *TO* is the acronym of training object. Equation 3.8 and Equation 3.4 have already provided the definition of *Texture* and *Compactness* feature. *Size* is simply the number of pixels of the object.

3.3.2.5 Link the Sub-objects of the Target Object (Step 5)

The proposed approach operates on the features of the training object and its sub-objects. Hence, the features of all the sub-objects are aggregated to represent a single feature for the operation with the training object. For aggregation within eCognition™, a built-in function *create links* of eCognition™ is employed. The function needs a common property of all the sub-objects for linking with each other. To provide a common property, all the sub-objects are manually classified into the land cover class of the training object. The linked objects are used to calculate the statistics, such as the sub-object with maximum size and size-weighted average of the *Texture* feature values of all the sub-objects.

3.3.2.6 Define the Scale, Shape, and Compactness Equations (Step 6)

The approach uses the *Texture*, *Compactness*, *Spectral stability*, and *Size* features of both the target object and its sub-objects to define the equations of the three parameters: Scale, Shape weight, and Compactness weight. The estimation equations of these three parameters are described next.

3.3.2.6.1 Scale parameter estimation

The scale parameter is the most critical parameter because it decides the average size of the resulting segments. Moreover, square of the scale parameter is the merging threshold (or heterogeneity threshold) of the segmentation process (see Equation 3.7 and Section 3.2.1). Therefore, the estimation equation of the scale parameter should use the size information of the objects. In addition, the major contribution to the equation should be from the spectral features [Benz *et al.*, 2004; Definiens AG, 2009].

Taking into account these two requirements, this thesis defines the scale parameter as:

$$Scale\ (S) = Current\ scale + \left(1 - w_{shape}\right) * \sqrt{(TextureTO - TextureSO} + \sqrt{(SizeTO - max\_sub\_size\ )}\ , \tag{3.11}$$

where *Current scale* is the scale parameter of the current level of segmentation at a particular iteration; w_{shape} is the shape weight; and *max_sub_size* is the maximum size values among the sizes of the sub-objects; and finally, *TextureSO* is size-weighted average of *Texture* feature values of all the sub-objects. This thesis defines *TextureSO* as:

$$TextureSO = \frac{1}{n_{merge}}.\sum_{i=1}^{t} n_i.Texture_i\ , \tag{3.12}$$

where t is total number of sub-objects; n_{merge} is the total number of pixels of the training object; and $Texture_i$ is the *Texture* feature value (see Equation 3.8) of sub-object i.

In Equation 3.11, the shape weight term minimizes the spectral heterogeneity (see Equation 3.1 and 3.7) when the shape weight is high. The *Current scale* represents the heterogeneity of the current segmentation level. *TextureSO* and *max_sub_size* features represent the heterogeneity of the current level of segmentation in the sub-objects. The subtractions of these two features from the target object features aim to remove the heterogeneity of the sub-objects from the target objects. This is to avoid redundancy of the heterogeneity, which is already added with the value of *Current scale*. *TextureSO* and *max_sub_size* features are derived by linking the sub-objects (see Step 5 of the workflow).

3.3.2.6.2 Shape weight estimation

The shape weight parameter can be viewed as a way to reduce the contribution of spectral heterogeneity in Equation 3.7. Figure 3.4 demonstrate the idea with the sub-objects of a building land cover training object. In the Figure, the small size sub-object (encircled in yellow) avoids merging with its adjacent large size sub-object because of the high spectral difference between the two sub-objects. For merging the sub-objects, there are two options:

1. Increase the scale parameter without increasing the shape weight: To allow the required increase in the spectral heterogeneity for the merging; and
2. Increase both the scale parameter and the shape weight: To reduce the effect of spectral heterogeneity on the overall heterogeneity (see Equation 3.7). This results in less increment of scale value as compared to the last option.

The second option minimizes the overall heterogeneity because the increment in the scale parameter would be less. Hence, the second option should be favoured [Tian and Chen, 2007].

In simple words, the shape weight should aid in the merging where the spectral heterogeneity restricts the merging. However, the sub-objects of Figure 3.4 may not be the sub-objects of the training object at lower segmentation level. These sub-objects may arise during the iterative process of multiresolution segmentation related to higher segmentation level (see Section 3.2). Nevertheless, the best estimate from the sub-objects of the training object can be used to simulate the situation of sub-objects of Figure 3.4.

To simulate the situation, the sub-object of the training object with the highest size-multiplied *Spectral stability* feature value is identified. The feature value of this sub-object with respect to the mean of the feature values of all sub-objects is related to the shape weight. For example, if the highest stability value is much higher than the mean value, then the shape weight should be increased. This thesis defines the relationship of this increase as the ratio:

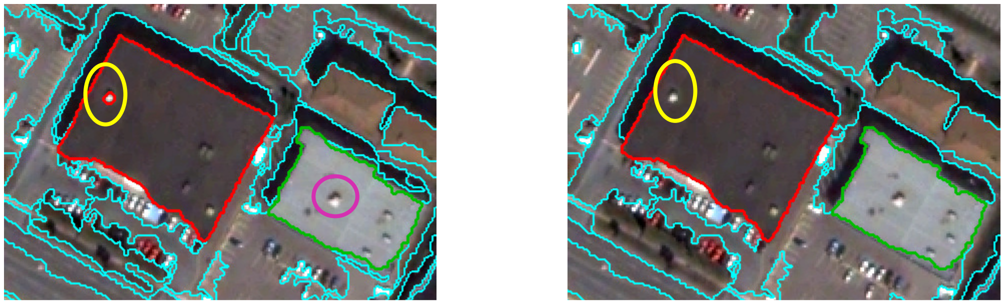
$$\frac{w_{shape}}{1-w_{shape}} = \frac{maximum(Spectral\ stability * Size\ of\ sub\ objects\ within\ x\ std\ of\ Y_i)}{mean(Spectral\ stability * Size\ of\ t\ sub\ objects)}\ , \tag{3.13}$$

where *std* refers to the standard deviation of Y_i. Y_i is the brightness difference with respect to average brightness of all sub-objects. Y_i is defined as:

$$Y_i = brightness_i - 1/t \,.\sum_i^t brightness_i\ , \tag{3.14}$$

x is a user-defined parameter (usually considered to be 1) to be multiplied with *std*.

Table 3.1 shows the calculation associated with the determination of maximum of *Spectral stability *Size* using Y_i of Equation 3.1.3. Y_i is selected instead of normal brightness to reduce the effect of size and increase the effect of spectral difference in the ratio. In simple words, Equation 3.1 represents the ratio of shape and spectral weight of the Equation 3.7. The required shape weight can be obtained by a simple rearrangement of Equation 3.13. However, the rearrangement would be less intuitive towards the logic of the increase of the shape weight. Hence, this paper defines the shape weight in this manner.



[a] A small size sub-object (encircle in yellow and white in color) has high spectral difference with respect to its adjacent sub-object (red) of the building. Hence, it is not merged with the adjacent sub-object. However, for the case of the building object (in green) the small size sub-object (encircled in pink) is merged since the spectral difference is not high.

[b] It shows how a slight increase in scale 100 to 105 with high increase in shape value 0.1 to 0.6, helps in merging the small size object with high spectral difference, as encircled in yellow.

Figure 3.4: Illustrates how spectral differences among the two adjacent objects affect the merging in multiresolution segmentation (MS). The parameters of MS in [a] is Scale = 100, Shape = 0.1 and Compactness = 0.5 and in [b] it is 105, 0.6 and 0.8.

Table 3.1: Demonstrates the calculation associated with the selection of sub-objects satisfying the condition to be considered for the determination of maximum Spectral*Stability with respect to Y_i of the numerator of Equation 3.13.

Sub-object no.	1	2	3	4	5	Statistics
Brightness(B)	10	35	45	42	38	Average(B) = 34
Y_i	-24	1	11	8	4	$Std(Y_i) = 13.95$ $Max(Y_i) = 11$
x	1	2	3			
$Y_i - x*13.95$	-2.95	-16.89	Only sub-objects with the Y_i value greater than the -2.95 will be considered for the maximum calculation, where x=1. The value -16.89 is considered for x = 2			

3.3.2.6.3 Compactness weight estimation

The compactness weight is linearly dependent on the *Compactness* feature value of the training object, i.e., if the *Compactness* feature value of the target object is high, then the weight should be high and vice-versa. In raster analysis, the most compact object is square so the minimum value of *Compactness* feature is 4 (see Equation 3.3) and the corresponding compactness weight is 1. The maximum value is arbitrarily chosen as 22 for a highly non-compact object. This value corresponds to the compactness weight 0. With these values, this paper defines the linear equation of the compactness weight as:

$$w_{compactness} = -0.056 * CompactnessTO + 1.1 \tag{3.15}$$

Ideally, the compactness weight evaluation should involve the smoothness feature value. However, the smoothness heterogeneity is always 4 times less than the compactness heterogeneity (see appendix I for the proof). Hence, the contribution of smoothness is ignored in the compactness weight evaluation.

3.3.2.7 Perform Segmentation with the Estimated Parameter (Step 7)

Using the parameters evaluated in the step 6, the results of the next/higher level of segmentation are generated. The segmentation result of the training object of this step is used for the comparison with the reference segment of the training object.

3.3.2.8 Test for Convergence (Step 8)

If the segmentation results of the training object converge to the reference segment of the training object, then the parameters are accepted. The criterion for the convergence is based on the visual assessment. In the assessment, the desired border of the training object is compared with the border produced from the segmentation results. If the visual comparison is successful, then the process of the workflow goes to step 9. Otherwise, steps from 3 to 8 are repeated for the next iteration. The iteration goes on until the results converge to the desired boundary or under-segmented result of the training object is achieved. The acceptance of the under-segmented results depends on the operator and the application of the segmentation results.

3.3.2.9 Evaluate Accuracy with respect to the Reference Objects (Step 9)

The reference objects for the accuracy assessment can be an external vector boundary layer or a vector boundary layer created within eCognition^{TM}. Then, a multiresolution segmentation is performed exclusively based on this vector layer [Definiens AG, 2009]. The results of the above segmentation process provide the reference outline of the objects. Using the built-in "*create link*" function of eCognition^{TM}, the extent of areal overlap among the reference objects and the segmentation results of Step 7 are evaluated

The extent of overlap is user-defined and 80 percent overlap is used for this paper. These overlaps also determine the required global accuracy of segmentation.

3.3.2.10 The End (Step 10)

This step marks the termination of the workflow with the accepted parameters and the global accuracy results of the final segmentation level. However, if the accuracy results are unacceptable, then the operator can change the training object and start the process of the workflow again. The condition of acceptance can be user-dependent or based on the percentage of global accuracy.

3.4 Experiments and Analysis of Results

A total of four experiments are performed using the parameters of the proposed approach in eCognition™ Developer 8.0. In the experiments, the estimated segmentation parameters for different land cover classes (Trees, Grass, and Buildings) are used for the segmentation of the images containing the land cover classes. Then, the segmentation results from the parameters of both the proposed approach and the trial and error based approach (see Section 3.1.2) are compared with each other. The comparison is based on a global segmentation evaluation technique proposed in this thesis. The next few sub-

Sections describe the experimental data sets, the estimated parameters, the final segmentation results, and analysis of the results in terms of their accuracies.

3.4.1 Image Data Sets

The image datasets for the experiments consist of three images of different sensors and locations: (1) a pan-sharpened Quickbird image of Fredericton city; (2) a pan-sharpened IKONOS image of Fredericton city; and (3) a pan-sharpened GeoEye-1 image of Hobart city in the island state of Tasmania, Australia. The images with four multispectral bands (blue, green, red, and near infra-red) were pansharpened using UNB-PanSharp software [Zhang, 2004]. Each of the pan-sharpened images has 11 bits of radiometric resolution and different spatial resolutions: 0.7 m for Quickbird, 1.0 m for IKONOS, and 0.5 m for GeoEye-1. The purpose of pan-sharpening is to enhance the spatial details of the images before their segmentation.

The major feature classes of the Quickbird image of Fredericton area are: (1) the commercial buildings, (2) residential buildings, (3) roads, (4) trees, (5) grasslands, (6) small patches of bare lands, (7) parking lots, and (8) urban forests. As shown in Figures 3.5[a] and [b], the classes of the Hobart area image and the Fredericton area image are the same. However, the number of commercial buildings in the Hobart scene is less compared to Fredericton scenes. The images have a good variety of different land cover objects, location, and sensors. This variety is essential to establish the effectiveness of the proposed approach.

[a] A subset of scene of Hobart, Australia scene from Geo-Eye-1 showing residential buildings.

[b] A subset of scene of Fredericton city, Canada from Quickbird showing commercial buildings.

Figure 3.5: Illustrates the classes of two images of different locations and sensors.

3.4.2 Experimental Results and Accuracy

The experiments utilized all of the four pan-sharpened multispectral bands of the images for the segmentation. All of the four experiments have been assigned with equal spectral layer weights (i.e., w_i's of Equation 3.1 are all equal to 1) for all the four bands of the images. The equal weights are acceptable results because these weights provide effective results in most of the applications of RS [Hofmann, 2001]. All the three images were subjected to an initial over-segmentation with the same initial parameters for the first three experiments. Table 3.2 shows the used default parameters for the initial over-segmentation.

In the fourth experiment, a multi-level analysis is performed to demonstrate the applicability of the proposed approach for the segmentation of hierarchically connected land cover classes. For example, the classes of small size Forest and large size Forest are hierarchically connected. Using the process defined in the workflow (see Figure 3.3), the final segmentation parameters were estimated and used for segmenting the images. Next few sub-Sections present the processing steps of the experiments and their results for the selected land cover feature classes using the proposed approach.

3.4.2.1 Segmentation of Residential Buildings of GeoEye-1 Hobart Scene

The Hobart city image has residential building objects with different roof colors as well as roof structures. Hence, Residential Buildings land cover class has high spectral and shape heterogeneity. However, the sizes of the building objects class of the image have similar scale of observation. Hence, a single scale of segmentation should be able to segment the building objects of the image.

With this background knowledge about the image, the steps of the experiment and their results for the Residential Building class are mentioned below:

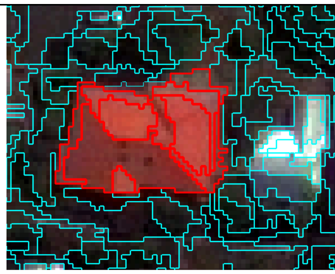
1. Use default parameters for the initial over-segmentation (Figure 3.6[a]);
2. Train the approach using a training object of the Residential Building class (Figure 3.6[a] and 3.6[b]);

3. Define the customized object features within eCognition^{TM} (Figure 3.6[c]);
4. Link the sub-object using built-in "*create links*" function of eCognition^{TM} (Figure 3.6[d]) for sub-object statistics;
5. Obtain the parameters using the Equations 3.11 to 3.15 from eCognition^{TM} object information list (Figure 3.6[c] and 3.6[e]).
6. Generate the segmentation results with the estimated parameters (Figure 3.6[f] and 3.6[g]).

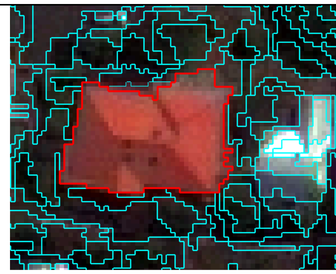
The characteristics of the estimated parameters for the Residential Buildings class are as follows:

1. Shape weight is high because the rooftops have high spectral heterogeneity. The heterogeneity is because of the inclination of rooftops with respect to sunlight.
2. Compactness weight is high because of the compact shape of the training object.

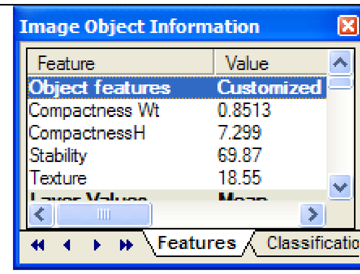
The above-mentioned steps are only for the parameter estimation approach in this thesis. In addition, a second set of three parameters are estimated using the traditional trial and error based approach (see Section 3.1.2). The traditional approach also resulted in the same delineation as shown in Figure 3.6 [f]. The parameters of the experiment resulted from both the approaches are shown in Table 3.2.



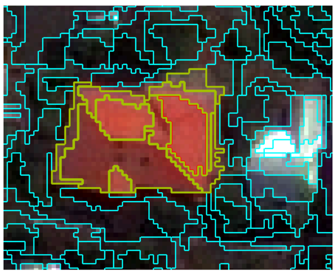
[a] Initial over-segmentation and sub-objects of a training object (red).



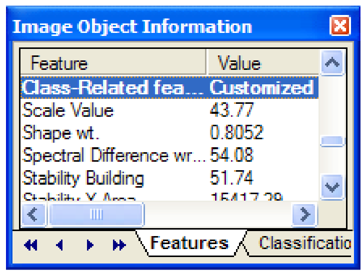
[b] Merge sub-objects to get the training object for supervised training.



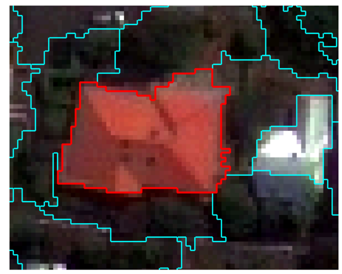
[c] Define compactness, spectral stability, and texture feature within eCognition[TM].



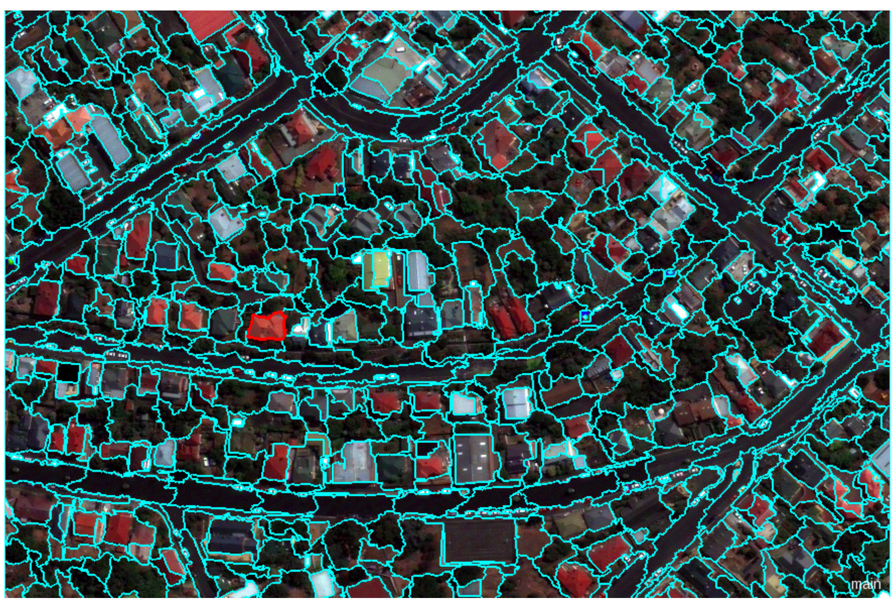
[d] Linked sub-objects statistics calculation (yellow).



[e] Scale parameter and Shape weight parameter evaluation.



[f] The training object segmentation obtained with the estimated parameters.



[g] Segmentation results using the parameters estimated by the proposed approach on other residential buildings of the Hobart scene.

Figure 3.6: Illustrates the general steps of the experiment of segmentation of residential buildings of Hobart city area.

Table 3.2: Segmentation parameters derived based on the proposed workflow as well as the trial and error method for Residential Buildings land cover class segmentation of the Hobart city image.

Parameter	Initial Over-segmentation	Final solution in 1 iteration	Trial and Error
Scale	20	43.77	34
Shape	0.1	0.8052	0.8
Compactness	0.5	0.8516	0.8

3.4.2.2 Segmentation of Grass Lands of Quickbird Fredericton Scene

The grass lands of Fredericton uptown area are homogeneous. The training object of the Grass Lands land cover class is selected as a stadium with grassy area and elliptical shape (Figure 3.7[b]). Figure 7 illustrates the experiment conducted to detect the grassy stadium with the results of: (1) initial over-segmentation (Figure 3.7[a]); (2) training of the approach using the training object (Figure 3.7[a] and [b]); 3) the training area segments results with the estimated parameters (Figure 3.7[c]); and (4) the segmentation results with the estimated parameters (Figure 3.7[d]).

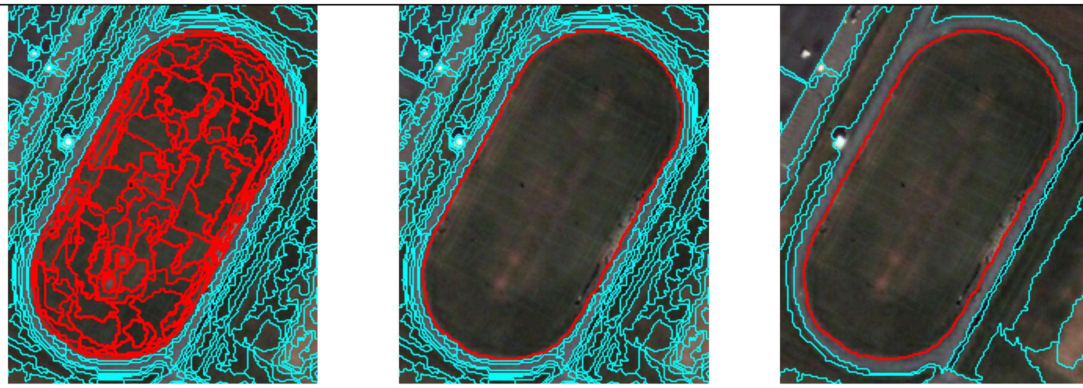
The estimated parameters have following properties:

1. High scale value compared to the scale value of the Buildings class. This is because of the large size of grass objects as compared to the building objects.
2. Low shape weight due to spectral homogeneity of the Grass Lands class.

Similar to the Buildings class, the trial and error approach resulted in the same delineation of the training object as by the proposed approach. The obtained parameters of the experiment from both the approaches are shown in Table 3.3.

Table 3.3: Segmentation parameters in each iteration of the proposed workflow as well as the trial and error method for grassy lands segmentation of Quickbird image of Fredericton city

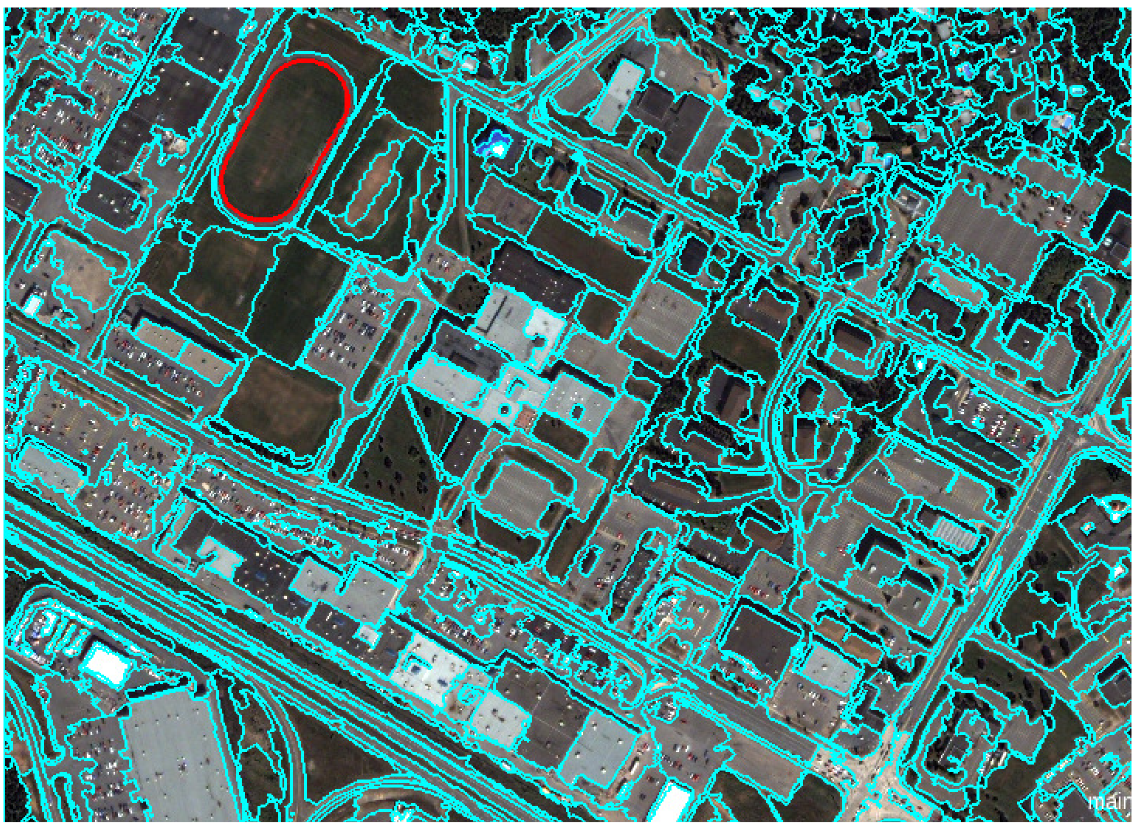
Parameter	Initial	Iteration 1 (Final)	Trial and Error
Scale	20	163.55	120
Shape	0.1	0.324	0.6
Compactness	0.5	0.8355	0.8



[a] Initial over-segmentation and sub-objects of the grassy stadium (in red) as training object.

[b] Merged sub-objects to get the training object for the supervised training.

[c] The training object segmentation obtained with the estimated parameters.



[d] Segmentation results using the parameters estimated by the proposed approach on other grass lands areas of Quickbird image of Fredericton area.

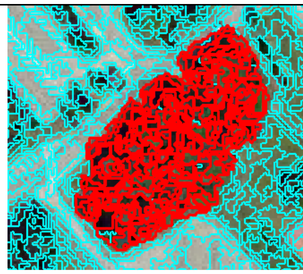
Figure 3.7: Illustrates the training and the results of the experiment of segmentation of grassy stadium of Quickbird image of Fredericton area.

3.4.2.3 Segmentation of Urban Forests of IKONOS Fredericton Scene

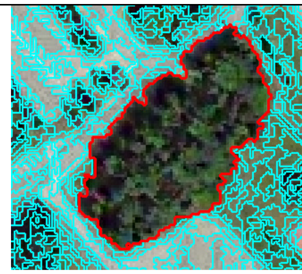
This section describes the multiscale capability of the proposed approach by the experiment on the Urban Forest Land cover class. The Urban Forests class of the Fredericton area is heterogeneous and textured. However, the class has different shapes and sizes. Hence, the Forest class has different scales of observation. Thus, an appropriate delineation of the forest patches requires a multi-level/multi-scale segmentation [Benz et al., 2004].

In the experiment, at first, the small size urban forests are identified. Figure 3.8 illustrate the results of the experiment involving small size urban forests: (1) the initial over-segmentation (Figure 3.8[a]); (2) the selected training object of the small size Urban Forests class (Figure 3.8[b]); (3) the results of segmentation using the estimated parameters on the training object (Figure 3.8[d]); and d) the results of the segmentation on the whole area of image (Figure 3.8[d]).

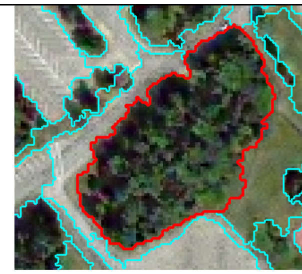
Unlike the Buildings and Grass classes, the segment delineating the training object shows a slight under-segmentation (Figure 3.8[c]). The under-segmentation result in the above experiment occurs because the contrast of the training object with its background is low. Nevertheless, the extent of overlap with the area of the training object is more than 90%. Hence, the segmentation results are acceptable.



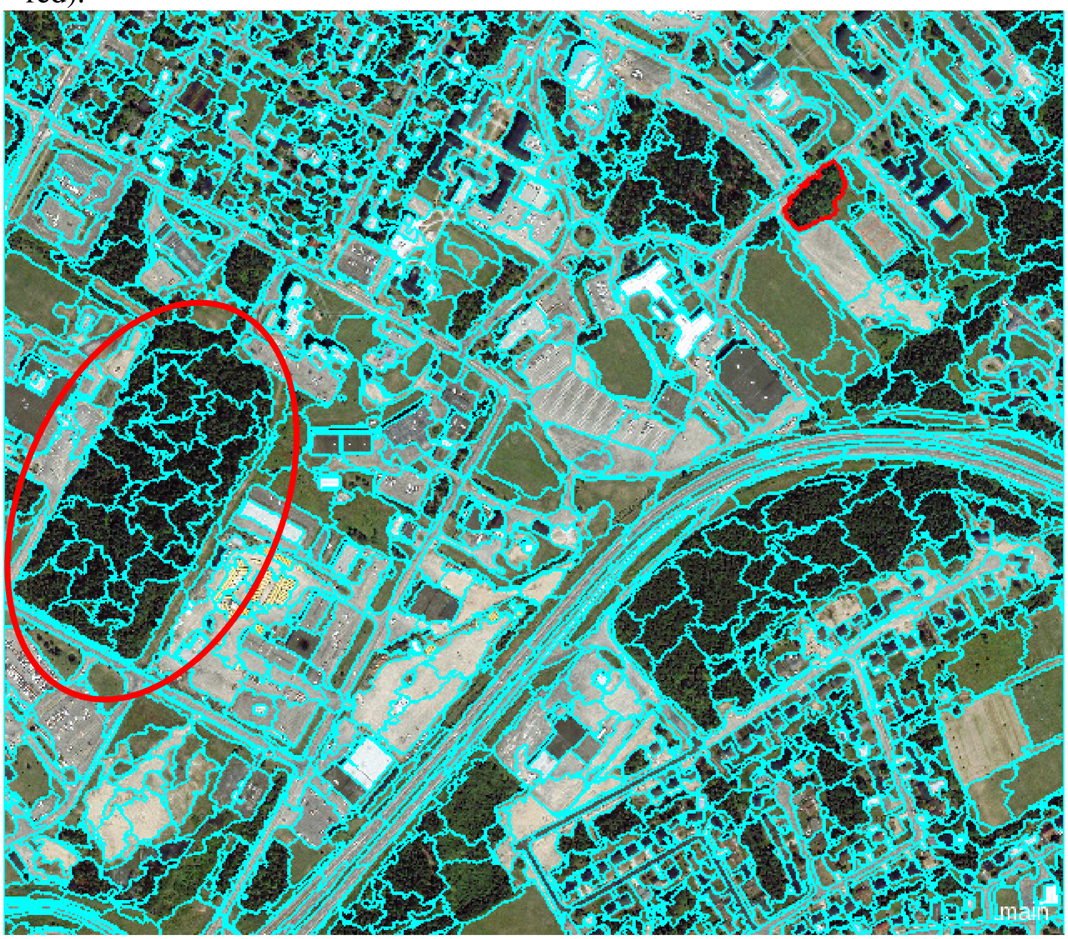
[a] Initial over-segmentation and sub-objects of the small size urban forest (in red).



b) Merged sub-objects to get the training object for supervised training



c) The training object segmentation obtained with the estimated parameters.



(d) Segmentation results using the parameters estimated by the proposed approach. The large size urban forest (in red) is over-segmented.

Figure 3.8: Illustrates the training and the results of the experiment of segmentation of small size urban forests of IKONOS image of Fredericton area.

After detection of small size urban forests, the large size forest objects are still over-segmented (see Figure 3.8[d]). Hence, the current segmentation level (shown in Figure 3.8[d]) can act as an initial over-segmentation for the segmentation of large size forest objects. Using this idea, the experiment was continued to segment large size urban forest objects at higher scale of segmentation. Figure 3.9 illustrates the results of the continued experiment: (1) the over-segmentation results on the large size forest objects Figure 3.9[a]); (2) the training object delineation (Figure 3.9[b]); (3) the results of the accepted segmentation using the estimated parameters (Figure 3.9[c]); and (4) the segmentation results on other large size urban forest objects (Figure 3.9[d]).

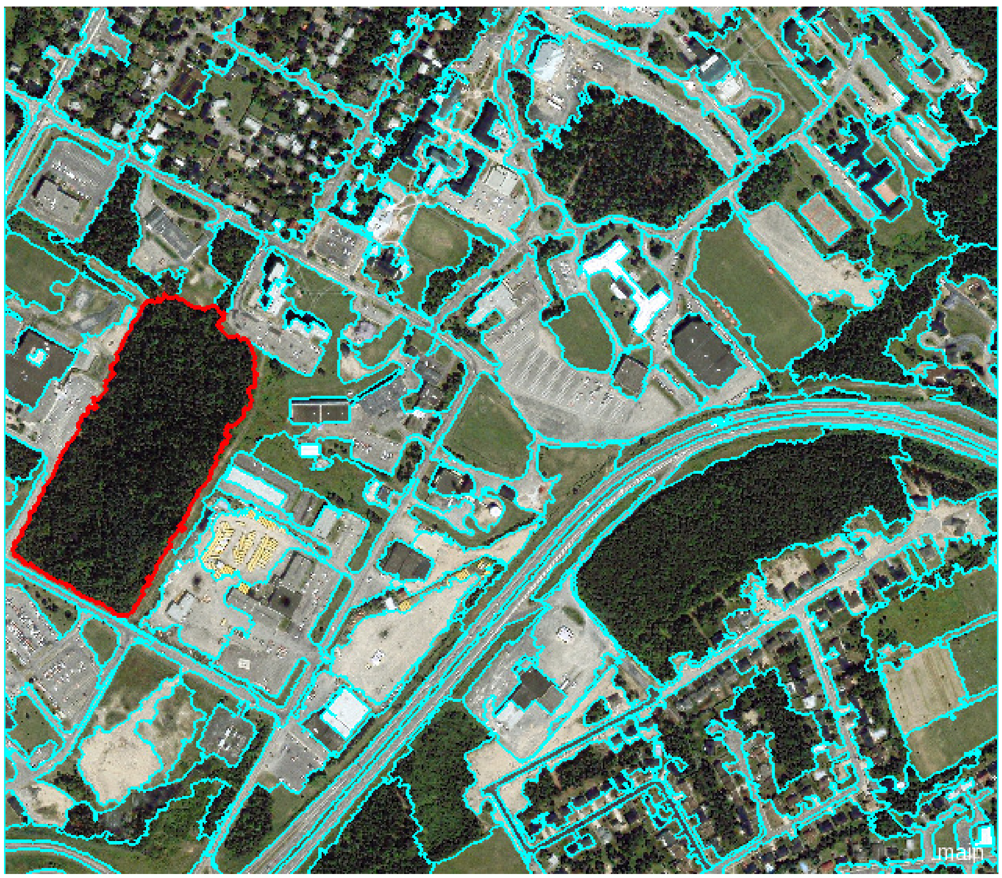
The shape parameters of both the small and large urban forest objects are high due to high spectral heterogeneity of the Urban Forest classes. The compactness parameters are also high because of the compact shapes of the training objects of the classes. However, the major difference lies with the scale parameter. The difference also demonstrates the size dependency of the scale parameter.

The trial and error approach also resulted in the under-segmentation of the training object of small size Urban Forest class. Moreover, the under-segmented result was different from the under-segmented result obtained from the proposed approach. However, the result of trial and error approach was also acceptable due to more than 90% overlap with the area of the training object. The final parameters used for segmenting the different scales of the urban forest objects using both the approaches are shown in Table 3.4 and Table 3.5.

(a) Sub-objects of urban forest of large size with initial over-segmentation (in red).

[b] Merged sub-objects to get the target object for supervised training.

[c] The training object segmentation obtained with the estimated parameters.



[d] Segmentation results using the parameters estimated by the proposed approach for large size forest segmentation.

Figure 3.9: Illustrates the training and results of the experiment of segmentation of urban forest of large size of IKONOS image of Fredericton area.

Table 3.4: Estimated Segmentation parameters from the proposed approach and the trial and error approach for large size of Urban Forest land cover class using IKONOS image of Fredericton city.

Parameter	Initial Over-segmentation	Final parameter in 1 iteration	Trial and Error (Small size)
Scale	20	75.45	120
Shape	0.1	0.8418	0.8
Compactness	0.5	0.7727	0.8

Table 3.5: Estimated Segmentation parameters from the proposed approach and the trial and error approach for large size of Urban Forest land cover class using IKONOS image of Fredericton city.

Parameter	Initial Over-segmentation	Final Parameters in 1 iteration	Trial and Error (Large size)
Scale	75.45	327.01	220
Shape	0.8418	0.7652	0.8
Compactness	0.7727	0.7464	0.7

3.4.2.4 Accuracy Assessment

The traditional procedure of obtaining the effective parameters is based on trial and error. Hence, the proposed approach based segmentation is compared with the trial and error approach based segmentation. However, the comparisons require a reference vector layer for each of the land cover classes used for the experiments. Hence, reference vector boundary layers of the land cover class are created manually using features of eCognition™.

The reference layers of the each of the land cover classes are compared with the two different segmentation results: 1) obtained using the proposed approach and 2) obtained using the trial and error approach. The parameters of the segmentation using the proposed approach and the trial and error approach are listed in Tables 3.2 to 3.5. The final step of the accuracy assessment is the global accuracy evaluation using the built-in "*create link*" function of eCognition™.

The "*create link*" function links the objects based on the extent of overlap between the reference objects and the obtained objects (see Figure 3.10[a]). The obtained objects can be the segments obtained after the segmentation with the estimated parameters from either the traditional or the proposed approach. The extent of overlap is determined by the user. Based on the number of links, the amounts of over-segmentation and under-segmentation are determined. The measures of over-segmentation and under-segmentation are described next.

The obtained objects, which satisfy the linking conditions of specified amount of overlap with respect to area of the reference objects, are categorized as the objects without over-segmentation (OWO). However, the current procedure fails to detect the under-segmentation (see Figure 3.10[b]). Therefore, the *create link* function is used again. However, in the second use, the overlaps are calculated with respect to area of the obtained objects instead of the area of the reference objects. The obtained objects, linked in this second usage of *create links* function, are categorized as the objects without under-segmentation (OWU). However, OWU include the over-segmented objects. Hence, the number of objects which exist in both the OWO and OWU are considered to be appropriately delineated objects as per the specified amount of overlap. Tables 3.6, 3.7, 3.8, and 3.9 enlist the OWO, OWU, the number of appropriately delineated objects, and number of samples used in the accuracy assessment for the land cover classes.

[a] Overlap area of two objects (grey). The percentage can be calculated using object 1 as well as object2.

[b] Under-segmentation area (grey) of object 2 with 100% overlap with the reference object (object 1in white).

Figure 3.10: [a] shows the overlap calculation; and [b] and how the overlap fails to detect under-segmentation.

Table 3.6: Accuracy results of the segmentation of the residential buildings of Hobart city using the parameters estimated from the proposed approach as well as the trial and error procedure.

Parameter estimation method	OWO (80% overlap)	OWU (80% overlap)	Appropriately Delineated (OWU∩OWO)	Total Buildings of Reference Layer	Accuracy (%)
Proposed	62	37	27	90	30
Trial and Error	46	45	27	90	30

Table 3.7: Accuracy results of the segmentation of the grassy lands of Quickbird image of Fredericton city using the parameters estimated from the proposed approach as well as the trial and error procedure.

Parameter estimation approach	OWO (80% overlap)	OWU (80% overlap)	Appropriately Delineated (OWU∩OWO)	Total Grasslands of Reference Layer	Accuracy (%)
Proposed	15	15	12	20	60
Trial and Error	16	17	14	20	70

Table 3.8: Accuracy results of the segmentation of the small size urban forest objects of IKONOS image of Fredericton city using the parameters estimated from the proposed approach as well as the trial and error procedure.

Parameter estimation approach	OWO (80% overlap)	OWU (80% overlap)	Appropriately Delineated (OWU∩OWO)	Total small urban forests of Reference Layer	Accuracy (%)
Proposed	12	10	10	14	72
Trial and Error	10	11	8	14	60

Table 3.9: Accuracy results of the segmentation of the large size urban forest objects of IKONOS image of Fredericton city using the parameters estimated from the proposed approach as well as the trial and error procedure.

Parameter estimation approach	OWO (80% overlap)	OWU (80% overlap)	Appropriately Delineated (OWU∩OWO)	Total large urban forests of Reference Layer	Accuracy (%)
The proposed	3	3	3	4	75
Trial and error	3	3	3	4	75

3.4.3 Analysis of Results

As per Table 3.5 through 3.8, the differences in the number of appropriately delineated objects by the trial and error and the proposed approach are low for all the four land cover classes. The accuracy results (Table 3.6 and 3.9) of the Residential Buildings class and the large size Urban Forests class are exactly same for trial and error and the proposed approach. However, the proposed approach has better accuracy in the case of small size Urban Forest class whereas the trial and error has better accuracy for the Grass Lands class.

The results based on the accuracy assessment are also dependent on the selection of reference vector boundary layers and spectral heterogeneity of the land cover objects. For example, the residential buildings have very low percentage accuracies using both the approaches due to the high spectral and shape heterogeneity of the residential buildings land cover class. On the other hand, the reference layers affect the case of the large size Grass Lands class and the large size Urban Forest class. The affect of reference layers is because of the difficulty in an appropriate manual delineation of the ambiguous boundaries of the large size objects. Hence, the results for the large size objects are likely to change since different operators will select different boundaries for the reference layers [Lang et al., 2009].

The second criterion of the accuracy is based on the visual assessment. Based on the visual assessments, the proposed approach based segmentation of the Residential Buildings land cover class has relatively higher under-segmentation as compared to the trial and error based approach. However, both the approaches provide the same accuracy result based on number of appropriately delineated objects (see Table 3.5). This result might be because the under-segmentation using the proposed approach is compensated by the over-segmentation using the trial and error approach. The segmentations results of the rest of land cover classes have very less visual differences. Hence, as per the visual assessment, the results of the trial and error and the proposed approaches are comparable with each other.

The final analysis is based on the parameters obtained by the proposed and trial and error based approaches. The shape weight parameter of the proposed approach increases with the increase in the heterogeneity of the feature classes and vice-versa. For example, high shape weights are obtained for the spectrally heterogeneous Urban Forest classes and low for the spectrally homogeneous Grass Lands class. This demonstrates that the proposed approach is adaptive with respect to the spectral properties of the training objects.

The proposed approach estimated high scale values for the large size objects as compared to the trial and error approach. However, the higher scale values have not affected the accuracy results (see Tables 3.7 and 3.9). This suggests that the effective parameters of the multiresolution segmentation are non-unique. This fact is also justified by Tian and Chen [2007] and Platt and Rapoza [2008].

3.5 Conclusions and Recommendations

This chapter proposed a semi-automatic supervised approach to estimate the three parameters (scale, shape weight, and compactness weight) of multiresolution segmentation. The parameters have been estimated using the available functionalities of eCognition™. Further, a global accuracy evaluation approach has also been proposed for comparing the segmentation results. The proposed approach scores over the traditional trial and error approach parameter estimation and other approaches because one or more of the following reasons:

1. **Fast**: The proposed approach provides the estimated parameters within 5 minutes for all the land cover classes. The limit of 5 minutes is irrespective of land cover classes and their locations. However, the same is not true for other solutions (see Section 3.1.3). Table 3.10 provides a time comparison of the proposed approach with other existing approaches including trial and error based approach. The duration of other approaches are as either reported by the research paper which proposed the approach or by the analysis of the authors.
2. **Effective**: Based on both the qualitative and qualitative segmentation evaluation techniques, the results are similar and comparable to the trusted trial and error based approach.
3. **Supervised approach**: It incorporates the knowledge of land cover classes for customized and multi-level estimation.
4. **Easy to implement**: The proposed approach is the easiest to implement because it uses features and functions of eCognition^{TM}.

Overall, the major benefit of this approach is its **fast performance and ease in implementation**.

Table 3.10: Comparison of the proposed approach with the other solutions of multiresolution segmentation with respect to time, implementation requirements, and solution assessment techniques.

The Solution	Implementation	Time	Assessment
Maxwell [2005]	Fuzzy logic (External application required)	Variable: 30 min to 1 hr; Source: Zhang et al. [2010]	Visual
Möller et al., 2007	Based on object metrics (External application required)	Not specified	Objective
Tian and Chen [2007]	Definition of meaningful segmentation (External application required)	Variable depends on number of parameter settings	Objective
Costa et al. [2008]	Genetic algorithms (External application required)	Variable (Generally in hrs)	Visual and objective
Marpu et al. [2010]	Definition of meaningful segmentation (External application required)	Variable depends on number of parameter settings	Objective
Drăgut et al. [2010]	Auto-correlation (External application required)	Variable (Dependent on image size and resolution)	Objective
Trial and Error	Based on idea of scale, shape and compactness stated in Definiens AG [2009]	Highly variable (Generally in hrs, see Zhang et al. [2010])	Visual
The Proposed Method	Object features (No External application required)	Always less than 5 minutes	Both visual and objective

Based on the accuracy results and the parameters (see Table 3.2 through 3.8), it can be concluded that the effective parameters for the multiresolution segmentation of a land cover class are not unique [Tian and Chen, 2007; Platt and Rapoza, 2008]. Hence, an operator's supervision is desired for the segmentation. The proposed approach successfully achieved the objective of assisting an operator's judgement because of the use of the supervised methodology and a global segmentation evaluation technique.

The future research with the proposed approach involves experimentation on low resolution satellite images (Landsat TM 7) as well as for a multi-level segmentation with land cover classes of different scales of observation. Moreover, the applicability of the results of the proposed in feature extraction applications, such as building detection and road extraction, can be experimented.

ACKNOWLEDGEMENT

This research is funded by the Strategic Project Grants Program of the Natural Science and Engineering Research Council of Canada (NSERC). The authors also like to thank city of Fredericton and ISPRS for the images used in this research.

REFERENCES

Baatz, M. and A. Schape (2000). "Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation." *Angewandte Geographische Informations-Verarbeitung XII*. Ed. J. Strobl et al. AGIT Symposium, Salzburg, Germany, 2000, pp. 12-23. [Internet], cited 12 Aug. 20011. Available in .pdf format through www.ecognition.cc/download/baatz_schaepe.pdf.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I. and M. Heynen (2004). "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information." *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 58, No.3-4, pp. 239-258.

Blaschke, T. and J. Strobl, (2001). "What's wrong with pixels? Some recent developments interfacing remote sensing and GIS." *GIS – Zeitschrift für Geoinformationssysteme,* Vol. 6, No. 1, pp. 12-17.

Blaschke, T., 2010, Object based image analysis for remote sensing. *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 65, No. 1, pp. 2-16.

Blaschke, T., C. Burnett, and A. Pekkarinen (2006). "Image segmentation methods for object based analysis and classification." In *Remote Sensing Image Analysis: Including the Spatial Domain,* Ed. Jong, S.M. de and Meer F. D. van der. Springer-Verlag, Dordrecht, The Netherlands, pp. 211-236.

Burnett, C., and T. Blaschke (2003). "A multi-scale segmentation/object relationship modelling methodology for landscape analysis." *Ecological Modelling*, Vol. 168, No. 3, pp. 233-249.

Castilla, G., and G. J. Hay (2008). "Image objects and geographic objects." In *Object-Based Image Analysis,* Eds., Blaschke, T., S. Lang and G.J. Hay, Springer-Verlag, Berlin, Germany, pp. 679-695.

Castilla, G., G. J., Hay, and J. R. Ruiz-Gallardo (2008). "Size-constrained region merging (SCRM): An automated delineation tool for assisted photointerpretation." *Photogrammteric Engineering & Remote Sensing*, Vol. 74, No. 4, pp. 409-419.

Chant, T. D., and M. Kelly (2009). "Individual object change detection for monitoring the impact of a forest pathogen on a hardwood forest." *Photogrammetric Engineering & Remote Sensing,* Vol. 75, No. 8, pp. 1005-1023.

Chen, J., D. Pan, and Z. Mao (2009). "Image-object detectable in multiscale analysis on high-resolution remotely sensed imagery." *International Journal of Remote Sensing*, Vol. 30, No. 14, pp. 3585-3602.

Costa, G.A.O.P., R. Q., Feitosa, T. B. Cazes, and B. Feijó, (2008). "Genetic adaptation of segmentation parameters." In *Object-Based Image Analysis,* Blaschke T., S. Lang and G.J. Hay (Eds.), Springer-Verlag, Berlin: Hiedelberg, Germany, pp. 679-695.

Drăgut, L., D. Tiede, and S. R. Levick (2010). "ESP: a tool to estimate scale parameter for multiresolution image segmentation of remotely sensed data." *International Journal of Geographic Information Science,* Vol. 24, No. 6, pp. 859-871.

Definiens AG, 2009. eCognition Developer 8.0, Reference Book, published by Definiens AG. Germany.

Flanders, D., M. Hall-Beyer, and J. Pereverzoff (2003). “Preliminary Evaluation of eCognition Object-Based software for Cut Block Delinieation and Feature Extraction.” *Canadian Journal of Remote Sensing*, Vol. 29, No. 4, pp. 441-452.

Haralick, R.M. and L. G. Shapiro (1985). “Image segmentation techniques.” *Computer Vision, Graphics, & Image Processing,* Vol. 29, No. 1, pp. 100-132.

Hay, G.J., T. Blaschke, , D.J. Marceau, and A. Bouchard (2003). “A comparison of three image-object methods for the multiscale analysis of landscape structure.” *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 57, No. 5-6, pp. 327-345.

Huang, H., B. Wu, and J. Fan (2003). “Analysis to the Relationship of Classification Accuracy, Segmentation Scale, Image Resolution.” In *Proceedings of IGARSS 2003 IEEE. Institute of Electrical and Electronic Engineers Inc*., 21-25 July 2003, Toulouse, France, , pp. 3671-3673.

Hofmann, P., (2001). “Detecting urban features from IKONOS data using object-oriented approach.” In *Proceedings of the First Annual conference of the Remote Sensing and Photogrammetry Society,* 12-14 September, RSPS, Munich, Germany, pp. 79-91.

Kim, M., M. Madden, and T. Warner (2008). “Estimation of optimal image object size for the segmentation of forest stands with multispectral IKONOS imagery.” In *Object-Based Image Analysis,* Ed. T. Blaschke, S. Lang and G.J. Hay. Springer-Verlag, Berlin: Hiedelberg, Germany, pp. 291–307.

Lang, S., E. Schöpfer, and T. Langanke (2009). “Combined object-based classification and manual interpretation-synergies for a quantitative assessment of parcels and biotopes.” *Geocarto International*, Vol. 24, No. 2, pp. 99-114.

Lavigne, D.A., G. Hong, and Y. Zhang (2006). “Performance Assessment of Automatic Feature Extraction Tools on High Resolution Imagery.” In *ASPRS MAPPS/ASPRS 2006 Fall Conference*, November 6-10, 2006, San Antonio, Texas.

Lizarazo, I., and J. Barros, 2010. Fuzzy Image Segmentation for Urban Land-Cover Classification, Photogrammetric Engineering & Remote Sensing, Vol. 76, No. 2, pp. 151-162.

Marpu, P.R., M. Neubert, H. Herold, and I. Niemeyer (2010). “Enhanced evaluation of image segmentation results.” *Journal of Spatial Science,* Vol. 55, No. 1, pp. 55-68.

Maxwell, T., 2005. Object-Oriented Classification: Classification of Pan-Sharpened QuickBird Imagery and a Fuzzy Approach to Improving Image Segmentation Efficiency. M.Sc.E. thesis, Department of Geodesy and Geomatics Engineering Technical report No. 233, University of New Brunswick, Fredericton, New Brunswick, 157 pp.

Möller, M., L. Lymburner, and M. Volk (2007). "The comparison index: A tool for assessing the accuracy of image segmentation." *International Journal of Applied Earth Observation and Geoinformation*, Vol. 9, No. 3, pp. 311–321.

Neubert, M., H. Herold and G. Meinel, G. (2008). "Assessing image segmentation quality-concepts, methods and applications." In: *Object Based Image Analysis*, Eds. Blaschke T., S. Lang and G.J. Hay, Springer-Verlag, Berlin, Germany, pp. 760–784.

Platt, R. V., and L. Rapoza (2008). "An evaluation of an object-oriented paradigm for land use/land cover classification." *The Professional Geographer*, Vol. 60, No.1, pp. 87-100.

Reed, R.T. and J.M.H. Du Buf (1993). "A review of recent texture segmentation and feature extraction techniques." *Computer Vision Graphics and Image Processing: Image Understanding,* Vol. 57, No.3, pp. 359-372.

Shackelford, A.K. and C.H. Davis (2003). "A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 9, pp. 1920-1932.

Smith, G.M., and R.D. Morton (2010). "Real World Objects in GEOBIA through the Exploitation of Existing Digital Cartography and Image Segmentation." *Photogrammetric Engineering & Remote Sensing*, Vol. 76, No. 2, pp. 163–171.

Tian, J., and D. M. Chen (2007). "Optimization in multi-scale segmentation of high-resolution satellite images for artificial feature recognition." *International Journal of Remote Sensing,* Vol. 28, No. 20, pp. 4625-4644.

Wuest, B., and Y. Zhang (2009). "Region Based Segmentation of QuickBird Multispectral Imagery through Band Ratios and Fuzzy Comparison." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 64, No. 1, pp.55-64.

Zhang, Y.J. (1997). "Evaluation and comparison of different segmentation algorithms." *Pattern Recognition Letters,* Vol. 18, No. 10, pp. 963-974.

Zhang, Y., T. Maxwell, H. Tong, and V. Dey, (2010). "Development of a supervised software tool for automated determination of optimal segmentation parameters for eCognition." *Proceedings of ISPRS TC VII Symposium- 100 Years ISPRS*, 5-7 July, Vienna, Austria, pp. 690-696.

Zhang, Y. (2004). Highlight Article: "Understanding Image Fusion." *Photogrammetric Engineering & Remote Sensing,* Vol. 70, No. 6, pp. 675-661.

CHAPTER 4

BUILDING DETECTION USING MULTI-LEVEL SEGMENTATION WITH A FUZZY PARAMETER BASED REGION MERGING CRITERIA

This chapter contains a conference paper published in an international conference and referred as:

Dey, V., Y. Zhang, and M. Zhong, and B. Salehi (2011). "Building Detection using Multi-Level Segmentation with a Fuzzy Parameter based Region Merging Criteria." Proceedings of the *32nd Canadian Symposium on Remote Sensing,* Sherbrooke, Québec, Canada, 13-16 July 2011, pp. 1-8.

The goal of the presented paper is to analyse the applicability of multiresolution segmented results (generated using a parameter estimation approach) for building detection, one of the basic requirement of all urban applications [Mesev, 2003]. The paper uses the parameter estimation approach of Maxwell [2005] instead of the approach proposed in Chapter 3. This is because: (1) the approach of Chapter 3 and Maxwell [2005] have very similar segmentation results; (2) the approach of Maxwell [2005] is more established than the proposed approach [Zhang et al., 2010]; and (3) the goal of the paper is to verify the plausibility of the idea of optimal parameter estimation with respect to an application. In order to present clearly, the published paper is slightly edited.

Abstract

Satellite imagery has broadened its fields of applications because of the increase in the spatial resolution of satellite images. Few such applications include urban planning, traffic monitoring, urban growth and management. Most of the urban applications require efficient building detection. This research aims to extract building objects from a pan-sharpened very high spatial resolution GeoEye-1 satellite image of Hobart, Australia. At first, a subset of image is selected which has building objects of approximately same scale of observation. This selection gives a rough estimate of the order of size of buildings in the image. In order to extract preliminary building sub-objects, a lower level multi-resolution segmentation is performed. Hue image and Sobel edge image are used along with RGBNIR bands for lower level segmentation. The generated segmentation results have improved edge extraction and roof-edge suppression compared to the segmentation results using only the pan-sharpened RGBNIR bands. After segmentation, shadow contexts of the buildings are used to identify the buildings objects. This identification is followed by a region-based Gaussian maximum likelihood classification for non-building objects. This classification is utilized to remove any falsely-detected building objects. The extraction performance is evaluated using both qualitative and quantitative analysis. The results are promising and with proper modifications might be used for real applications.

4.1 Introduction

Enhancement of spatial resolution in optical remote sensing has been phenomenal since the successful launch of IKONOS in 1999. IKONOS provides 1m panchromatic (Pan) and 4m multispectral (MS) imagery. Few more examples of currently available very high resolution (VHR) satellites are as follows: (1) Quickbird (0.7m Pan and 2.8m MS) launched in October 2001, and (2) GeoEye-1 (0.41m Pan and 1.65m MS) launched in September 2008. These satellites have enlarged the scope of applications of satellite images and encroached in the field of applications of aerial images, such as transportation planning, city development planning, urban planning, object change detection, urban monitoring, land use and land cover map development, and GIS database update and management [Ünsalan and Boyer, 2005].

4.1.1 Need of Building Detection using VHR Satellite Imagery

Building detection is one of the fundamental feature extraction tasks in various urban applications [Song et al., 2006]. Normally, aerial images are used for the building detection because aerial images can provide 3D height information from DEM. Although most of the developed urban cities have aerial images, they might be outdated considering the rapid urban growth [Ünsalan and Boyer, 2005]. Therefore, the building detection results from the aerial images might not be up-to date.

On the other hand, satellite images are widely available as well as accessible with wider coverage as compared to aerial images and LIDAR data [Konecny and Schiewe, 1996]. Moreover, urban applications are also possible using VHR Satellite images (Spatial resolution <= 1m). Further, a human interpreter delineates the building boundaries from 2D VHR images with almost 100% certainty. Hence, building detection using VHR imagery seems to be a viable option to obtain up-to-date information of the locations of buildings.

4.1.2 Building Detection Approaches

As mentioned in the last Section, a human interpreter easily identifies the buildings from an urban VHR image. A human interpreter utilizes semantic and contextual properties in order to extract buildings from the images. Some of these properties are as follows: (1) edge information, (2) context information like shadow; (3) prior knowledge, e.g., buildings near roadside and parking lot buildings, and (4) color homogeneity of rooftops. Thus, it is reasonable to use the properties utilized by a human interpreter in developing an automatic approach. However, modelling these properties possess a great challenge due to complex shapes and sizes of buildings in various locations [Mayer, 1999; Benedek et al., 2009]. Overall, the formulation of a computer-based (automatic) building detection approach is a non-trivial task but it is possible [Ünsalan and Boyer, 2005].

Researchers have proposed several approaches based on the: (1) edge and shadow context of buildings, (2) stochastic analysis, (3) statistical procedures, (4) snake model, and building structure model [Mayer, 1999; Peng et al., 2005; Mayunga et al., 2007; Benedek et al., 2009; Luan and Ye, 2010]. Further, researchers have also proposed building detection based on image segmentation (Song et al., 2006; Shackelford and Davis, 2003). This paper utilizes the image segmentation based procedure for building detection because it is simple and has no assumptions regarding the shape of the building. Moreover, image segmentation is one of the best automatic approaches to simulate the interpretation of a human.

4.1.3 The Proposed Approach and Objectives

This paper employs an image segmentation-based approach to derive the outlines of the buildings of a residential area. It selects multiresolution (MR) segmentation technique (implemented in commercial software eCognition™) for the segmentation. To enhance the building boundary delineation, edge image and hue image are used as additional inputs layers along with the pansharpened multispectral VHR image for the multiresolution segmentation [Zhang, 2004]. However, MR segmentation suffers from parameter estimation problem [Schiewe et al., 2001]. Hence, this paper utilizes fuzzy based supervised parameter optimization (FbSP) method for parameter estimation of MR segmentation [Maxwell and Zhang, 2005]. After MR segmentation, shadow context is used for building identification along with the: (1) prior knowledge of land cover classes, (2) size, (3) color tone from the hue image, and (4) geometric features of the segments. Finally, false building detection is minimized using a segment-based Gaussian maximum likelihood (GML) classification.

The overall objective of this research work is to study the effectiveness of FbSP in the context of feature extraction, i.e., building detection. In this paper, efficiency of the proposed detection approach is identified based on the number of correct boundary delineation and number of false detection [Lin and Nevatia, 1998]. The assumptions of the research of this paper are as follows:

1. Buildings can be detected as closed polygons.
2. Building roofs comprise of different shades of same color with different intensities. The shades are separable using the hue image, which is generated from HSI transform of the multispectral image of this study.
3. Scale of observation of the buildings on the VHR image is approximately same.
4. Buildings' cast detectable shadows.

The shadow context assumption is valid because most of remote sensing satellites are sun-synchronous and they take images during 10 to 11 am. During this time of the day, the sun's elevation angle is not vertical. Further, the angle is sufficiently inclined to cast shadow during this time.

The next two Sections describe the study area and methodology used for the experiments conducted to achieve the above-stated objectives. These two Sections are followed by the analysis of the obtained results and the conclusions of this research.

4.2 Study Area

The study area of this research is GeoEye-1 imagery of Hobart Australia. The image has four multispectral bands of 2m spatial resolution (Red, Green, Blue, and NIR) and one panchromatic band of 0.5 m resolution. A subset of image has been taken to include the residential buildings of same scale of observation [Dare, 2005]. Multispectral bands of the residential subsets are pansharpened using UNB pansharpening method, which is available in PCI Geomatica 10.0. The final spatial resolution of the pan-sharpened images is 0.5 m.

The buildings on the image are of different shapes, sizes, and colors. However, the rooftops are relatively homogeneous with different illuminations due to the inclination of rooftops with respect to sunlight. Such a scenario is true for most of North American and Australian residential buildings images but it is not for the residential buildings of

European countries [Luan and Ye, 2010]. Hence, the assumption of homogeneous rooftops is valid in the case of images of North America and Australia. Apart from the buildings, other classes of VHR image are roads, vegetation, bare land, and parking lots. Figure 4.1 shows a snapshot of the image, which is used for the experiments. The study image has a total of 559 buildings identified manually using polygon-based digitization.

Figure 4.1: The study area of GeoEye-1 imagery of Hobart, Australia having residential buildings of similar scale of observation.

4.3 Methodology

4.3.1 Multiresolution Image Segmentation

The first step of the proposed process is MR segmentation of the image. MR image segmentation requires three effective parameters for efficient results. To solve the problem of MR segmentation, this study utilizes the FbSP method, which was proposed by Maxwell and Zhang [2005]. The FbSP method requires a training object of a land cover class. This training object is formed by merging its sub-objects, which are formed by initial over-segmented image of the training object. The FbSP method utilizes the spectral, shape, and size properties of the training object and its sub-objects for parameter estimation. To summarize, FbSP essentially maps the properties of the target object and its sub-objects to determine the scale, shape, and compactness parameters [Maxwell and Zhang, 2005].

In FbSP, fuzzy logic is used via Fuzzy inference systems (FISs). These FISs utilizes five properties of training object and its sub-objects for effective parameter estimation. These properties are spectral standard deviation, spectral mean difference, size, compactness, and rectangularity. The resulting scale value aims to generate the objects with the scale to the target object. However, the method is inefficient in retaining edges of buildings and also suffers from the over-segmentation of the residential buildings due to different illuminations of building rooftops (shown in Figure 4.2[a]).

In order to retain the edges, Sobel edge gradient image is utilized as an additional layer for the multiresolution segmentation of the image. The hue image from HSI transformation is also used as an additional layer to suppress the building roof edge and the roof's illumination differences. This suppression is aimed to identify the building rooftops as a single segment. Figure 4.2[a] and [b] show the final results of MR segmentation with: (1) FbSP optimized parameters and without the additional layers and (2) FbSP optimized parameter with the two additional layers as mentioned above.

[a] [b]

Figure 4.2: Illustrates the benfit of using edge and Hue Image layer for segmentation. [a] shows the non-compact objects generated from normal MR segmentation without additional layers; and [b] shows the MR segmentation, with better results (yellow), using hue and edge image as additional layers.

4.3.2 Building Shadow Identification

Shadows can be extracted using simple spectral value thresholding [Dare, 2005]. While it is possible to detect shadows using simple thresholding, it does not necessarily give shadows of the required building regions. However, shadows are the major cues of building detection so it is cannot be avoided. Hence, this paper modifies the shadows detected results from simple thresholding for efficient shadow detection.

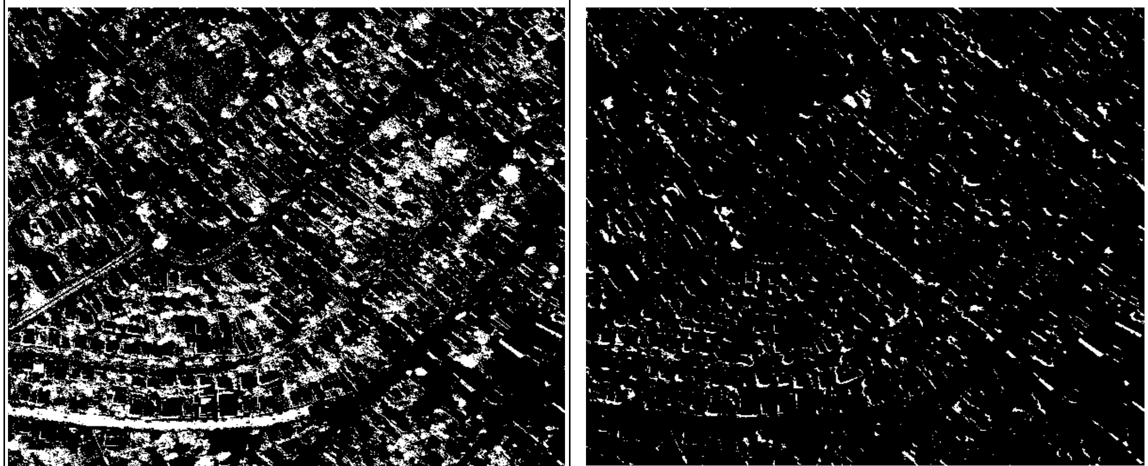
The initial shadows, which are detected using the simple thresholding, are mixed with (1) tree shadows, (2) black roads, (3) black cars, and (4) black rooftops. This study tackles only the first two types of the mixing. Regarding the third one (black cars), the spurious buildings detected by assuming black cars as shadows are eliminated at the post processing step. Finally, for the black rooftops, only one or two out of 559 buildings have the black rooftops. Hence, those shadows are ignored in the experiments. The following directions are employed for efficient shadow identification:

1. Use simple user-based thresholding based on the image histograms for initial shadow detection from the lower level of segmentation.
2. Identify connected regions of shadows using region growing.
3. Use morphological processing to remove shadows, which are falsely detected, e.g., shadow regions comprising of one or two pixels due to the shadows of cars.
4. Remove shadow regions which are higher than a specific size threshold. The average size of the objects is known because the objects are of

similar scales. Hence, the size of shadow cannot exceed the buildings size. Using this fact, a suitable threshold is selected.

5. Remove vegetation shadows by identifying the vegetation regions from the shadows using the shadow-object geometry, which is derived based on the sun's azimuth
6. Remove shadows with orientation angle less than 90 degrees. Since sun's azimuth is 38 degrees for the selected Hobart scene of GeoEye-1 image, the angles casted by the shadow regions should be greater than 90 degrees. The angles are identified based on the direction of the major axis of smallest fitting ellipse of the region with respect to pixel coordinate system.

Figure 4.3[a] and Figure 4.3[b] show the detected shadow images using simple thresholding and the final shadows obtained after the above-stated processing directions. On visual comparison, it can be said that the final shadows are more precise and accurate compared to the shadows detected using simple thresholding method.



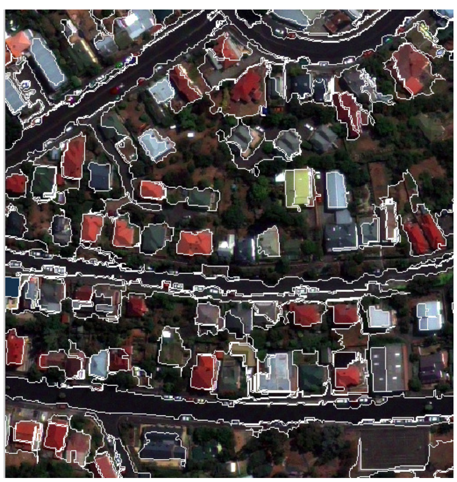
[a] [b]

Figure 4.3: Illustrates the shadow detection using the proposed approach. [a] shows shadows (with tree shadows, black roofs, and black roads) identified using simple thresholding approach; and [b] shows the shadow image after road segments and tree shadow removal using proposed approach.

4.3.3 Building Identification using Shadow-object Geometry

Buildings are detected from shadows using the shadow-object geometry. The geometry identifies the segmented regions along the sun's azimuth angle, where the starting points are centroid of the shadows [Sirmaçek and Ünsalan, 2008]. Figure 4.4[a] shows the identified buildings using the shadow-object geometry. The Figure shows lots of buildings with false detection. This false detection is mainly because of the tree shadows and black cars identified as shadows. These shadows detected roads and bare land objects as the building objects.

A segment-based (object-based) GML classification is performed to remove the false detection due to the objects of Bare Land and Roads class. The selection of bare land and roads class is known through the prior knowledge of adjacency of these classes with respect to the building objects. At first, statistics of the training samples of these two classes are utilized for a two class GML classification with 95% confidence value. Then, the building objects which are classified into these two classes are removed. After this step, the false building segments of vegetation class are removed using NDVI index. Figure 4.4[b] shows the residential buildings identified after the removal of the false building detection by GML classification.


[a]

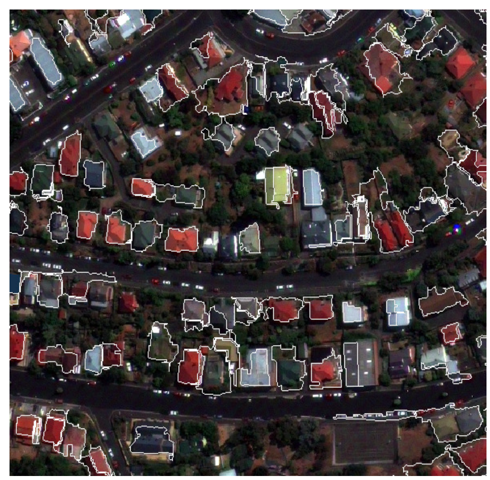

[b]

Figure 4.4: Illustrates the building detection with and without removal of the falsely detected buildings. [a] shows the buildings identified before GML classification: and [b] shows the identified buildings after GML classification, with removed spurious roads and bare land objects.

4.3.4 Accuracy Analysis

Building detection has a very standard accuracy assessment technique as proposed by Lin and Nevatia [1998]. It identifies both the false detection percentage and true detection accuracy based on the reference layer of the building objects. For this paper, the reference layer is building vector layer identified by manually digitizing the building outlines. For accuracy, true detection percentage of buildings, *DP,* is defined by Lin and Nevatia [1998] as:

$$DP = \frac{TP}{TP + TN} * 100 \tag{4.1}$$

where *TP* represents the true positive (number of buildings identified both manually and the automatic approach) and *TN* represents true negative (number of buildings identified manually but not by automatic approach).

On the other hand, the false detection percentage or the branching factor, *BF,* of buildings is defined by Lin and Nevatia [1998] as:

$$BF = \frac{FP}{TP + FP} * 100 \tag{4.2}$$

where *FP* represents the false positive (the number of buildings identified by the automatic approach but not in the reference layer).

For the identification of *TP,* two criteria are considered. These criteria are:

1. For the accuracy test, qualified building objects are those test which have the areal overlap of 80% and above with respect to the manually digitized 559 buildings.
2. Among the qualified buildings, the centroid difference between the detected building segments and the corresponding reference building outlines should not be more than 10% of the smaller side of the bounding box of the corresponding reference building. This criterion is to further filter the detected buildings based on the shift of the locations with respect to reference building outlines.

With these considerations, Table 4.1 summarizes the values obtained after the applied building detection method. The reported accuracy is not very high. However, considering the image complexity and the general assumptions the detected results are good.

Table 4.1: Summarizes the accuracy of the proposed building detection approach using multiresolution segmentation.

TP	TN	FP
400	103	56
DP	BF	
79.5%	12%	

4.4 Conclusions

Most of the building detection approaches assumes building shape to be a simple rectangular buildings, which may not work for complex-shaped (e.g., polygon) buildings [Mayer, 1999]. In this paper, a general building detection approach is proposed. The approach aims to detect the simple roof residential buildings of any shape. The approach utilizes a multi-level segmentation along with shadow context, scale of observation, prior knowledge, size, and spectral features for building detection.

The results show that the multiresolution segmentation with the FbSP estimated parameter produces results which can be applied for a general application, such as building detection application of this paper. Although the correct building detection is around 80%, the segmentation based building detection approach is plausible because the missed detection mainly correspond to the failure of shadow detection. Since shadows are the main cues for building detection, false or missed shadows would correspond to loss of detection of the buildings. Hence, the proposed building detection approach can be improved by improving the shadow detection.

Few suggested areas of further research are improvement in the shadow detection, customization of region merging after initial low level multiresolution segmentation, and incorporation of more of building objects' properties, such as detection of vertical wall of the building and the road proximity for the buildings.

ACKNOWLEDGEMENTS

This research is funded by the Strategic Project Grants Program of the Natural Science and Engineering Research Council of Canada (NSERC). The authors would also like to thank ISPRS website for free images, available for research purposes.

REFERENCES

Benedek, C., X. Descombes, & J. Zerubia (2009). "Building Extraction and Change Detection in Multitemporal Aerial and Satellite Images in a Joint Stochastic Approach." Research Report 7143, INRIA Sophia Antipolis, Rocquencourt, France, pp. 1-40. [Internet] Available online at: hal.archives-ouvertes.fr /docs/00/44/09/20/PDF/benedek_RR.pdf (accessed 11 Jan 2011).

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I. and M. Heynen (2004). "Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information." *ISPRS Journal of Photogrammetry & Remote Sensing*, Vol. 58, No.3-4, pp. 239-258.

Dare, P. M. (2005). "Shadow Analysis in High-Resolution Satellite Imagery of Urban Areas." *Photogrammetric Engineering & Remote Sensing*, Vol. 71, No. 2, pp. 169-177.

Konecny, G., and J. Schiewe (1996). "Mapping from digital satellite image data with special reference to MOMS-02." *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 51, No. 4, pp. 173–181.

Lin C., and R. Nevatia (1998). "Building detection and description from a single intensity image." *Computer Vision and Image Understanding*, Vol. 72, No.2, pp. 101–121.

Luan, Q., C. Ye, and J. Long (2010). "Review of building information extraction based on single high-resolution remote sensing images." In: *Proceedings of IEEE International Congress on Image and Signal Processing*. CISP2010, 16-18 Oct., Yantai, China, pp. 2245-2250.

Maxwell, T. L. and Y. Zhang (2005). "A fuzzy approach to supervised segmentation parameter selection for object-based classification." In Proceedings of SPIE: Applications of Digital Image Processing XXVIII, 2-4 Aug., San Diego, USA, Vol. 5909, pp. 528-538.

Mayer, H. (1999). "Automatic Object Extraction from Aerial Imagery- A survey Focusing on Buildings." *Computer Vision and Image Understanding*, Vol. 74, No. 2, pp. 138-149.

Mayunga, S.D., D. J. Coleman and Y. Zhang (2007). "A semi-automated approach for extracting buildings from QuickBird imagery applied to informal settlement mapping." *International Journal of Remote Sensing*, Vol. 28, No. 10, pp. 2343-2357.

Mesev, V., 2003, Remotely sensed cities: an introduction. In V. Mesev, Editor, *Remotely Sensed Cities,* pp. 1-19, (London, Taylor and Francis).

Peng J., D. Zhang and Y. Liu (2005). "An improved snake model for building detection from urban aerial images." *Pattern Recognition Letters*, Vol. 26, No. 5, pp. 587–595.

Schiewe, J., L. Tufte, and M. Ehlers, 2001. Potential and problems of multi-scale segmentation methods in remote sensing. *GIS-Geoinformationsysteme*, Vol. 6, No. 1, pp. 34-39.

Shackelford, A.K. and C.H. Davis (2003). "A hierarchical fuzzy classification approach for high-resolution multispectral data over urban areas." *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41, No. 9, pp. 1920-1932.

Sirmaçek, B., and C. Ünsalan (2008). "Building detection from aerial images using invariant color features and shadow information." In *IEEE International Symposium on Computer and Information Sciences*. ISCIS '08, 27-29 Oct., Istanbul, Turkey, pp. 1-5.

Song, Z., C. Pan, and Q. Yang (2006). "A region-based approach to building detection in densely build up high resolution satellite image." In: *Proceedings of IEEE International conference on Image Processing*. ICIP '06, 8-11 Oct., Taipei, Taiwan, pp. 3225-3228.

Ünsalan, C. & K. L. Boyer (2005). "A system to detect houses and residential street networks in multispectral satellite images." *Computer Vision and Image Understanding*, Vol. 98, No. 3, pp. 423-461.

Zhang, Y. (2004). "Highlight Article: Understanding Image Fusion." *Photogrammetric Engineering & Remote Sensing,* Vol. 70, No. 6, pp. 675-661.

CHAPTER 5

SUMMARY AND CONCLUSIONS

This chapter summarizes the research conducted in this thesis to achieve the goal of the improvement of segmentation results for the VHR satellite imagery. The chapter outlines the research and the contributions of Chapters 2 to 4 towards this goal. In addition, the chapter provides recommendations for the further research.

5.1 Summary of the Research

In the beginning of this thesis, it was identified that RS research has shifted from pixel-based analysis to object-based image analysis or OBIA to efficiently analyze the increased spatial details of VHR imagery [Blaschke, 2010]. OBIA required an efficient VHR image segmentation as the fundamental step. With a comprehensive review of research papers on segmentation (**Chapter 2**), multiresolution segmentation was found to be most widely-used segmentation technique for a VHR image analysis in OBIA. In the **Chapter 3**, it was found that the major problem of multiresolution segmentation is its parameter estimation. The traditional trial and error approach may take hours to estimate the efficient parameters. Hence, this thesis proposed a supervised parameter estimation approach which can estimate efficient parameters within 5 minutes irrespective of the sizes, locations, and types of land covers (**Chapter 3**).

The efficiency of the estimated parameters was proved by the successful experiments on different land covers classes: Residential Buildings, Grass Lands, and multi-scale Urban forests (see Figure 3.6 to 3.9). Moreover, a global accuracy evaluation technique has been proposed to establish the efficiency with respect to the performance with the trial and error approach (see Table 3.6 through 3.9).

In addition, the segmentation results with the estimated parameters are employed for a building detection application (**Chapter 4**). The building detection algorithm employed two additional layers (Hue image and Edge image) for multiresolution segmentation. With this modification, the building detection accuracy increased from 30% to 80% (see Table 3.6 and 4.1). This improvement justified the applicability of the proposed approach in general feature extraction problems.

5.2 Contributions of the Research

The **major contribution** of this thesis is the improvement of the performance of the parameter estimation of multiresolution segmentation. While other existing parameter estimation approaches may require hours to estimate the parameters, the proposed parameter estimation approach determines efficient parameters within **5 minutes for any land cover classes** (see Table 3.10). Moreover, the approach has been implemented within eCognition^{TM}. This implementation makes the analysis of the approach easier than other existing approaches. Hence, the approach is both fast and easy to analyze.

In order to validate the above achievement, a global accuracy evaluation technique was also proposed. This accuracy technique is also unique in the sense that the technique is completely implemented using features and functions of eCognition^{TM}. Hence, the proposed approach avoids the requirement of external software applications, i.e., software other than eCognition^{TM}. Overall, the proposed approach improves the performance of multiresolution segmentation by estimating efficient parameters within very short time (5 minutes).

The **other contributions** of this thesis include: 1) the identification of suitable categories of techniques and widely used techniques for VHR images/OBIA applications (**Chapter 2**) and 2) a building detection algorithm using the results of multiresolution segmentation (**Chapter 4**). In the review eight categories of the VHR image segmentation techniques were identified. These categories are as follows: (1) Clustering approach; (2) Level set model; (3) MRF model; (4) ANN model; (5) Fuzzy model; (6) Multi-scale model; (7) Watershed model; and (8) HSMR model. Out of these categories, multi-scale and watershed based techniques are identified as the most widely used. This identification is also a major contribution because there are hundreds of research papers available on image segmentation. Therefore, the proper guidelines proposed in this thesis for the selection of segmentation techniques are important for researchers, who are new to the segmentation field.

The building detection algorithm proposed in this thesis is more general in its assumptions (see Section 4.1.3). However, the accuracy of the detection is not too high (80% only). Hence, the applicability of the approach in real urban applications is doubtful.

Overall, in this thesis, the range of contributions includes: 1) identification of the suitable segmentation techniques, e.g., multiresolution segmentation; 2) improvement in the performance (in terms of duration) of parameter estimation for multiresolution segmentation; and 3) finally, application of the segmentation results from the parameter estimation approach for extraction of residential buildings.

5.3 Recommendations for Further Research

The research found that the watershed segmentation technique is a growing research field for VHR image segmentation. Hence, a comparison of performance of a popular watershed segmentation technique (e.g., SCRM by Castilla et al. [2008]) and the multiresolution segmentation would be a probable area of research.

Although the segmentation results using the solution to the problem of parameter estimation was justified for different land cover classes, a more comprehensive experimental set up is required to ascertain the performance on different land cover classes. Moreover, the solution produced very high scale parameter values (compared to trial and error for the same segmentation results) for the large size training objects. This is not desirable because the increase in scale parameter increases the heterogeneity. Hence, a potential research area lies with minimizing this increment in scale parameter value for the large size training object.

The shadow detection approach, utilized for the building detection, was found to be inappropriate. Hence, a more effective shadow detection approach is required to improve the building detection. Moreover, it was found based on the accuracy results that multiresolution segmentation might not be useful for building detection because of the complexity of the buildings' rooftops. Hence, a customized region merging (such as fuzzy region merging) for improvement in the building detection from urban VHR image is another area for further research.

REFERENCES

Blaschke, T., (2010). "Object based image analysis for remote sensing." *ISPRS Journal of Photogrammetry and Remote Sensing,* Vol. 65, No. 1, pp. 2-16.

Castilla, G., G. J. Hay, and J. R. Ruiz-Gallardo (2008). "Size-constrained Region Merging (SCRM): An Automated Delineation Tool for Assisted

Photointerpretation." *Photogrammetric Engineering & Remote Sensing*, Vol. 74, No. 4, pp. 409-419.

Appendix I

Mathematical Analysis of Equations of Multiresolution Segmentation

A1. Compactness Heterogeneity vs Smoothness Heterogeneity

The aim of this Section is to prove that the compactness heterogeneity is always at least four times greater than the smoothness heterogeneity. At first the definition of Compactness and Smoothness are re-represented from Equation 3.3 and Equation 3.4 as:

$$Compactness = \frac{l_{Obj}}{\sqrt{n_{Obj}}} \quad , \tag{A1.1}$$

$$Smoothness = \frac{l_{Obj}}{b_{obj}} \quad , \tag{A1.2}$$

The numerators of Equations A1.1 and A1.2 are same. Hence, the comparison should be for denominator. Now, b_{obj} represents perimeter smallest rectangle, which is greater than the area of the object enclosed by it. Suppose, length of b_{obj} is l and breadth is w, then area, A, is:

$$\mathrm{A} = l * w, \tag{A1.3}$$

and b_{obj} is:

$$b_{obj} = 2*(l + w)\ . \tag{A1.4}$$

As per the definition of the b_{obj}, the area of the rectangle is always less than the area of the object enclosed by it, represented as:

$$\mathrm{A} \geq\ n_{obj}. \tag{A1.5}$$

Using the Equation A1.3 and A1.4, it can be said that:

$$2 * (l + w) \geq 4 * \sqrt{l * w} \tag{A1.6}$$

From Inequality A1.5 and Equation A1.3, it can be concluded that:

$$b_{obj} \geq 4 * \sqrt{n_{obj}} \tag{A1.7}$$

Therefore, from Equation A1.1, A1.2, and Inequality A1.7 it can be concluded that compactness heterogeneity is always four times greater than smoothness heterogeneity.

Curriculum Vitae

Candidate's full name: Vivek Dey

Universities Attended:

Indian Institutes of Technology Kanpur, Kanpur, India, 2004-2008, B. Tech. in Civil Engineering.

Publications:

Dey, V., Y. Zhang, M. Zhong, and B. Salehi (2011). "Image Segmentation Techniques for Urban Land Cover Segmentation of VHR Images: Recent Developments and Future Prospects." *International Journal of Remote Sensing,* (under review).

Dey, V., Y. Zhang, and M. Zhong (2011). "A Supervised Methodology for Optimal Parameter Estimation of Multiresolution Segmentation within eCognition." *Journal of Photogrammteric Engineering and Remote Sensing,* (to be submitted).

Dey, V., Y. Zhang, and M. Zhong, and B. Salehi (2011). "Building Detection using Multi-Level Segmentation with a Fuzzy Parameter based Region Merging Criteria." Proceedings of the *32nd Canadian Symposium on Remote Sensing,* Sherbrooke, Québec, Canada, 13-16 July 2011, pp. 1-8.

Dey, V., Y. Zhang, and M. Zhong (2010). "A Review of image segmentation techniques with remote sensing perspective". Wagner W., and B. Székely (Eds.). *Proceedings of ISPRS TC VII Symposium- 100 Years ISPRS*, 5-7 July, Vienna, Austria, pp. 31-42.

Zhang, Y., T. Maxwell, H. Tong, and V. Dey, (2010). "Development of a supervised software tool for automated determination of optimal segmentation parameters for eCognition." *Proceedings of ISPRS TC VII Symposium- 100 Years ISPRS*, 5-7 July, Vienna, Austria, pp. 690-696.