Urban scene Description for a Multi scale classification of High Resolution imagery : case of Cape Town urban Scene

By Guy Blanchard Ikokou
Supervised by Dr George SITHOLE
Department of Geomatics

June ,2012

Thesis submitted to the University of Cape Town in partial fulfilment of the requirements for the degree of Masters of Science in Engineering.

Abstract

With the availability of high resolution data including aerial photographs, remote sensing applications have been tested in natural and man made environment such as urban areas. In South Africa, up to date urban information describing urban growth and development is missing in data banks at Municipal, and provincial levels. Information produced from remote sensing techniques such as image classification could eventually fill this gap. However, the quality of the produced information is dependant on the classification technique and strategy used. In this paper, a multi level contextual classification approach of the City of Cape Town, South Africa is presented. The methodology developed to identify the different objects using the multi level contextual technique comprised three important phases. The first phase involved the selection of high level features such as spectral, spatial, shape, size information, spatial distances. The second phase involved the selection of suitable segmentation parameters to apply the classification technique on and the third phase dealt with the contextual classification of land cover classes based on the urban scene model produced in the phase 1.

The results of this study revealed the effectiveness of the contextual classification approach. An overall accuracy and Kappa index of 88.069 and 83.700 percent were found. Classification accuracy of individual objects increased from the level 1 mostly pixel based, to the level 3 contextual classification. Certain land use classes have been accurately identified based on their descriptive contextual information. As examples, commercial, industrial, educational buildings, grass sport fields, recreation parks have been detected based on their size, shape, spectral characteristics and distance relations with a high accuracy.

Declaration

1. I know that plagiarism is wrong. Plagiarism is using another's work and pretend that it is my own.
2. I have used the Harvard convention for citation and referencing. Each significant contribution to and quotation in this project,from the work or works of other people has been attributed and has been cited and referenced.
3. This project is my own work
4. I have not allowed and will not allow anyone to copy my work with intention of passing it off as his or her own work.

Signed :

Guy Blanchard Ikokou

Acknowledgement

I would like to thank Dr George Sithole for his valuable supervision and support through the development and improvement of this thesis. This work would not have been completed without his guidance. Special thanks to my father, mother, brothers and sisters for the motivation through those difficult years in Cape Town. I would like to express my gratitude to the Geomatics Department for providing the data to be used in this study. I also would like to thank all my colleagues for their help and kindness to me. A special thank to my fiancée Minkoue Catherine for being next and supportive to me.

Contents

1 Introduction **1**
1.1 Background and context 3
1.2 Scope and limitations 3
1.3 Hypothesis 4
1.4 Aim and objectives 5
1.5 Thesis Outline 5

2 Literature review **7**
2.1 Investigation of some advanced image segmentation techniques 8
2.1.1 Region-based level set segmentation technique 8
2.1.2 Knowledge-based segmentation technique 9
2.1.3 Context-based segmentation approach 9
2.1.4 Texture-based segmentation techniques 11
2.1.5 Shape and Size based segmentation techniques 12
2.1.6 Neural network segmentation techniques 13
2.1.7 The Fuzzy clustering segmentation techniques 14
2.1.8 Temporal image segmentation techniques 16
2.1.9 Multi resolution segmentation techniques 16
2.2 An investigation of some advanced classification techniques . . 19
2.2.1 Pixel-based classification techniques 19
2.2.2 Sub-pixel classification techniques 20
2.2.3 Per-field classification approaches 21
2.2.4 Context-based classification techniques 23
2.2.5 Knowledge-based classification techniques 24
2.3 Investigation of some of the most recent scene description techniques 26
2.3.1 Size measurement techniques 27
2.3.2 Shape description techniques 27
2.3.3 Investigation of urban materials and their spectral properties 31
2.3.4 Measures of object context 34

2.4 Review of some urban classification studies 38
2.4.1 Investigation of road extraction studies 39
2.4.2 Urban water extraction techniques 40
2.4.3 Urban open space extraction techniques 41
2.4.4 Urban vegetation extraction techniques 41
2.5 Investigation of some post classification tools 42
2.6 Conclusion 42

3 urban scene contextual model for multi-scale classification 46
3.1 Introduction 46
3.2 Material and Methods 47
3.2.1 Study area and Data 47
3.2.2 Urban land use selection 50
3.2.3 Size and shape features 51
3.2.4 Extraction of objects spectral features 52
3.2.5 Geometric and Topological Relations 53
3.3 Experimental results 57
3.3.1 Variation in geometric relations 62
3.4 Discussion and conclusion 67

4 Optimal Scale Parameter selection for a Multi-scale Segmentation. 69
4.1 Introduction 69
4.2 Material and Methodical Approach 70
4.2.1 Material 70
4.2.2 Segmentation of images 72
4.2.3 Identification of suitable scale parameters 73
4.2.4 Assessment of scale parameters 74
4.3 Experimental Results 75
4.3.1 Identification of suitable scale parameters 75
4.3.2 Refining under and over segmentation 77
4.4 Discussion and conclusion 80

5 Multi level Image Classification 83
5.1 Introduction 83
5.2 Material and Methods 84
5.2.1 The data 84
5.3 Image segmentation 84
5.4 Multilevel contextual classification 86
5.5 Experimental results 93
5.6 Discussion and conclusion 100

6 Conclusion and Future work **102**
6.1 Conclusion 102
6.2 Future Work 104

List of Tables

3.1 Examples of high level spectral features considered in the investigation: on the left column are listed the spectral features and in the right column are listed their respective corresponding descriptions 54

3.2 Compactness values per land cover type: Roads and parking plots are examples of irregular shaped objects whereas buildings are examples of more regular objects. Similar shape signatures are observed between trees and water bodies on one hand and between Recreation areas and Parking plots on the other hand. 60

3.3 Subset of 11 samples of building patterns analysed using the nearest neighbour distribution index : Building pattern numbers 1, 9, 17, 23, 46 and 60 revealed indices closer to zero and were described as composed of clustered buildings. Building patterns numbers 3, 12,20, 39 and 91 indicated indices greater than 1 and were classified as composed of regularly structured building. Building patterns 23 and 60 had indices approximating 1 and were classified as having randomly arranged buildings. 64

3.4 Subset of high level samples derived from different residential scenes. Eight high level features are shown because of the large amount of data as the study was done on 92 residential subscenes. Scenes with building counts larger or equal to 30 were classified high density scenes. The letter M stands for mean in the table, MIBD stands for mean interbuild distance, MBRD stands for mean building-road distance and MSV stands for mean spectral value 66

4.1 A subset of the 1464 spectral and size samples used to compute internal segment homogeneity and spatial autocorrelation of Kensington urban area. 71

4.2 Weighted variances and Moran's indices produced by each investigated scale parameter. 76

5.1 Basic variable used in level three classification decision tree. The type column defines if the variable is spatial or spectral and the layer indicate the band from which the variable was extracted in case of spectral type 90
5.2 Error matrix of the level 3 classification. The word comm in the table stands for commercial buildings, Grass L for grassland class, Art. S.F for artificial sport fields, Recr.a. for recreation areas, O. space for open space, Educat for educational buildings, Indust for industrial buildings, Swim.p for swimming pools, Parking p for parking plots, Resident for residential buildings and other non-res for other non residential buildings 98

6.1 Classification statistics comparison between the three classification levels. Level 1 classification produced the lowest accuracy and Kappa index as it mostly relied on pixel information. 103

List of Figures

1.1 Research Design: After data projection the image was filtered before performing a scene analysis to derive information such as object topology, size and shape. A multi-scale segmentation was performed thereafter on the image before applying a multi-scale context-based classification to assign each object to its corresponding category. 2

2.1 An illustration of a data divided into four quadrants: A and C show large areas without objects and B and D show cluster organizations of objects. 36

3.1 The study site of Cape Town including the different suburbs falling into the area 48

3.2 After a visual analysis of the high resolution image, a set of land use classes is selected in order to derive the different high level features which are evaluated in order to build a scene model. 49

3.3 Figures A, B, C and D are three subsets of the study site illustrating the different land used samples selected in this investigation. Figure A shows an example of grassland, figure B shows examples of non residential buildings,figure C illustrates two examples of artificial sport fields and figure D reveals examples of parking areas, open space(bare ground), shrub and roads. 51

3.4 The polygon A on the left hand side image illustrates recreation parks, surrounded by trees and grassland. The two red polygons (B) on the top right image show examples of residential areas composed of small and medium size buildings of red, green and grey roofs. Polygon C on bottom right image illustrates a large dam(dark colour) surrounded by trees and buildings. 51

3.5 Spatial organization bar: Values closer to zero characterize clustered pattern,values closer to 1 describe random organization and values closer to 2.4 describe more regular patterns. . 55
3.6 The scene A shows a more regular spatial organization of buildings, the space between buildings is more or less constant. The scene B shows clustered building with high variation of inter building distances.These types of patterns generally score a nearest neighbour index of 1 as buildings are randomly distributed. 56
57
3.8 A and B :examples of Non residential buildings. From visual observation, the buildings have respective size larger than other surrounding buildings in the area. 58
3.9 Histogram plots showing trees and road dominant spectral ranges in the red (Layer 1) and green(Layer 2) bands: Trees, in dark black colour dominate the spectral range between 15 and 83.8 in the red band and 23 to 94 in the green band. . . . 60
3.10 The band contains relevant spectral information that enabled the separation between buildings and water, tress and grass sport fields. The remaining classes tend to overlap due to very small spectral distances. 60
3.11 The band holds significant spectral information describing trees,grassland and green roofs. Trees and buildings were separated by a quite good spectral distance of 1.0802. Grassland(sport fields) and trees were also separable with a distance of 1.670. Roads and building roofs exhibit similar spectral properties that made the separation difficult as the spectral distance between the two classes was 0.065. 61
3.12 Parking plots, grassland, recreation parks and artificial sport fields exhibited high spectral reflectance in the redness index and could be separated from water. Open space and buildings are examples of objects difficult to separate in this spectral index. 61
3.13 Roads and buildings can not be separated in this index with a spectral distance less than 1. Open space cannot be separated from buildings and trees because of the very small spectral distances that separate them. Building and trees exhibit similar spectral properties in this band as shown by the distance close to zero. 62

3.14 Recreation parks can be separated from trees in this band combination. Conversely, building roofs present high spectral similarities with roads, recreation parks and artificial sport fields, making their spectral separation very difficult during classification based on pixel values. 62
3.15 Sub-scenes from 14 to 87 are examples of high density areas with number of buildings per area greater than 30. Sub-scenes 4, 13 and 87 are examples of low residential areas with number of buildings per area less than 30. 63
3.16 Road-building distances: building patterns 4, 12, 24, 34, 36, 52 and 87 are examples of patterns located away from roads, whereas building patterns 16, 18, 28, 40 and 64 are examples of patterns located very close to roads. 63
3.17 Three examples of building spatial arrangement in Cape Town : A: tendency to random spatial organization,B: Tendency to linear clustering organization, and C: Tendency to more regular spatial organization. 65

4.1 A :Original Cape Town urban scene, B: Residential area of Kensington, composed of individual housing units of small and medium size, a large road network, trees and recreation areas. C: Residential area of Vredelhoek. The area also contains a large road network, grass land, artificial sport fields and trees. D: Oranjezicht residential area. 70
4.2 Overview of the methods used: Spectral and area measures are extracted from individual segments. These measures are then normalized and used to compute the objective function that identifies potential scale parameters. The potential scale parameters identified are then evaluated using an heterogeneity function. 72
4.3 Objective function identified seven low values suspected to be associated with suitable scale parameters but only six of them are related to negative Moran's index, characteristics of suitable scales (table 5.2). Scales of 15, 35 and 100 were found appropriate to segment Kensington urban scene, mostly dominated by small and medium size buildings; The scale parameter of 120 was found suitable to segment Vredelhoek area dominated by a large road network and the scales of 180 and 200 were found appropriate to segment forest extent in Oranjezicht area. 76

4.4 Heterogeneity function showing segments internal homogeneity at different segmentation scales. The higher the value of the index, the more homogeneous are the segments produced. Scales of 100 and 120 produced better segment homogeneity with higher values of the index. Scales of 35, 180 and 200 produced good enough segments homogeneity whereas scales of 25, 50,55, 65, 150, 260 produced poor results with high internal heterogeneity. 77
4.5 A: original Kensington scene, B: segmentation at scale of 10 showing building roofs partitioned into several smaller segments, C:segmentation at scale of 55 showing over and under-segmentation of the areas and D: segmentation at scale of 100 showing building roofs segments approximating their real world corresponding. 78
4.6 A: original Oranjezicht scene, B: segmentation at scale of 25 showing roads segmented into several smaller objects, C:segmentation at scale of 35 still showing small segments of roads and D: segmentation at scale of 120 showing road segments approximating their real world lengths. 79
4.7 A : the original Vredelhoek scene, B: segmentation results at scale of 25 showing individual trees, B: segmentation at scale of 55 still showing small trees segments and D: segmentation results at scale parameter of 120, individual trees have been grouped to form larger forest extents, suitable for classification. 80
4.8 A: segmentation results performed with the scale parameter of 150, the resulting objects contained more than one class including buildings, roads, open space and trees. B: segmentation done at scale parameter of 120, the segment comprises a portion of road that is optimally captured compared to results in A. 80

5.1 A: the type of the object cannot easily be determined due to insufficient contextual information. B: put into its real context, the object was identified as a building. This illustrates the difficulty encountered by computers to recognize unknown objects without context. 86
5.2 An example of a semantic network for the class building. At level one, buildings are observed as a town, at level two, buildings can be observed as residential and non-residential structures. At level three, buildings are perceived as single houses, block of flats, industrial and commercial structures. 87

5.3 Level 3 classification decision tree. Variables and corresponding thresholds are represented by the quadrilaterals. The different terminal nodes are represented by rectangles.The arrows indicate the direction of the decision which are represented by terminal nodes. As example an unknown object located at a distance smaller or equal to 30 meters from any building, with a length over width ratio ranging between 1 and 17 is likely to be a road segment. For a definition of the different variables see table 5.1 below. 89
5.4 A: segmentation results at scale parameter of 100 showing good delineation of small and medium size building units, B: image segmented at scale parameter of 120 showing good delineation of large size objects such as large buildings, grassland and parking lots ,C and D: image segmentation results at scale parameter of 180 showing good extraction of road network as well as group of trees. 93
5.5 Classification results at level 1 on the right hand side compared to the original image on the left hand side. Land covers were extracted based on pixel reflectance,size,shape and distance measures.The yellow circle shows a grassland accurately extracted based on pixel reflectance,size measure and shape compactness. 94
5.6 The classification at level 2, on the right hand side was improved by the use of context information that enabled the separation of residential from non residential buildings. The distance relation between buildings in classification level 1 on the left hand size and parking plots played an important role in distinguishing non residential from the building class as illustrated in the level 2 classification results. 95
5.7 Great improvements can be observed on the image on the right produced at level 3 classification, compared to the image on the left hand side produced at level 2. This improvement was possible by the use of relationships such as membership, border to, super object of,sub object of, used in this level. The red circles show classification improvements of a group of trees using the membership relation. 96

Chapter 1

Introduction

With the launch of new generations of high and very high spatial resolution imagery, remote sensing has garnered attention in the scientific community through civil and military applications. Most of remote sensing applications concentrate on classification of spatial objects located on the earth surface based on their spectral properties. However, this classification of spatial ob jects from aerial imagery remains a great challenge because the success of the process is influenced by many factors such as the characteristics of the scene, the type of remote sensing data used, and image processing techniques used. The main remote sensing systems providing high resolution data are the SPOT, IRS[1], Ikonos, Quick Bird, Geo-Eye, airborne multi-spectral scanners as well as digital aerial photography (Cordley, 1996; Ridley et al., 1997).

However, the commonly used automatic classification methods rely only on pixel values and use supervised and unsupervised algorithms which are suitable for low spatial resolution data. With very high spatial resolution data showing various land cover types, it is very challenging to obtain satisfactory results with supervised and unsupervised classification techniques. To overcome this limitation, object based image analysis methods have been tested .These approaches segment images into homogeneous regions in a hierarchical manner and produced image objects that can then be used for classification. The concept of hierarchy is a key element in object classification as research has shown the complexity of extracting object of various size at a single scale. Multi-scale context-based classification does not classify object based on their spectral properties alone but also considers other image objects properties such as shape, size characteristics, as well as its locale

[1]India Remote Sensing

and global context. The quality of derived objects, however, depends on the selection of segmentation parameters defined by the analyst.

This thesis proposes a multi-scale context-based classification technique for urban areas using high resolution images. The approach comprises four models : (1) The first model involves the pre-processing of the image including projection, image quality enhancement and image subset to select a smaller and representative study site in order to reduce computer processing time. (2) The second model deals with the selection of appropriate high level features to consider as inputs for image classification. This model includes shape, size, context, spatial relations and spectral signatures. The aim of selecting high level features is to build an urban scene contextual model that will be useful for urban land use applications as well as top down or bottom up image analysis and understanding. (3) The third model involves a technique of selection of optimal segmentation parameters to segment the study area and the last model deals with the multi-scale classification of the image. Figure 1.1 shows an overview of the methods that were used in this research:

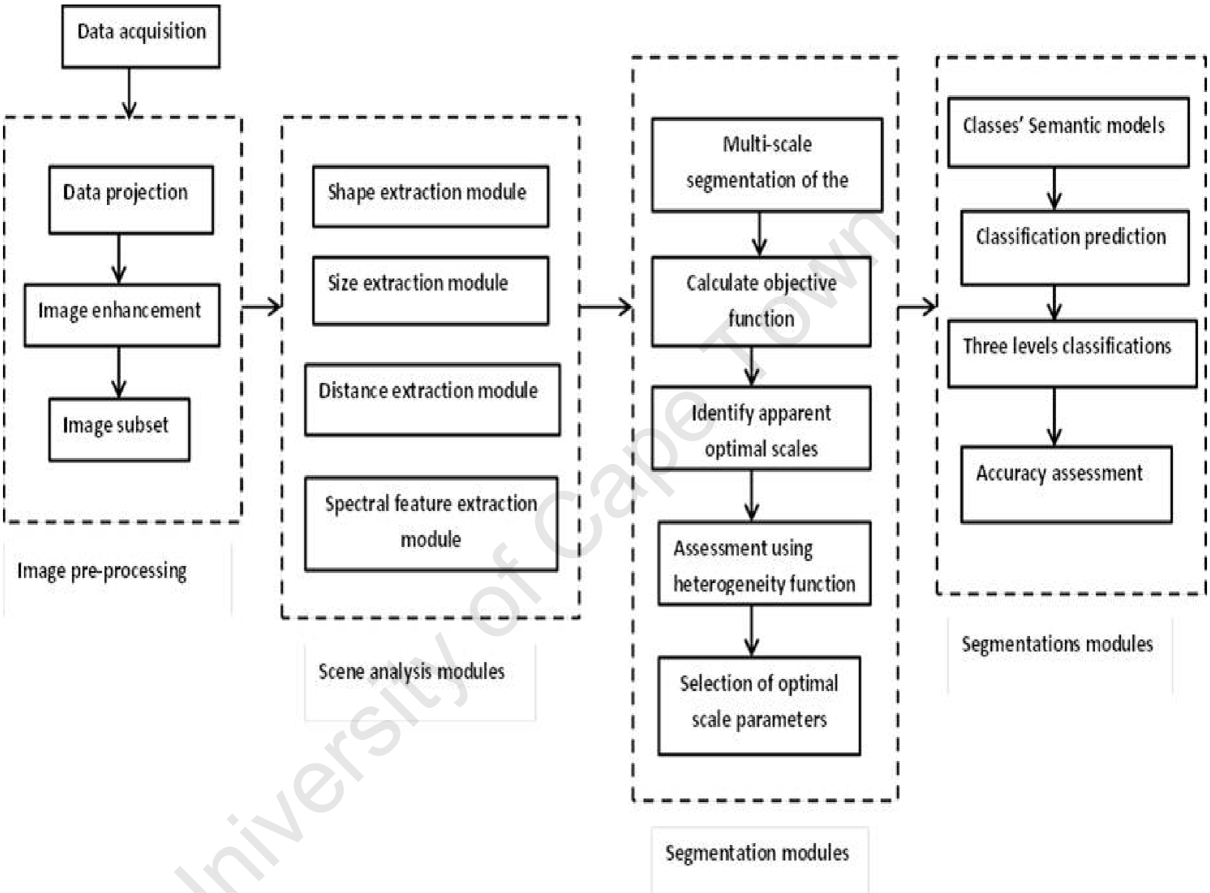

Figure 1.1: Research Design: After data projection the image was filtered before performing a scene analysis to derive information such as object topology, size and shape. A multi-scale segmentation was performed thereafter on the image before applying a multi-scale context-based classification to assign each object to its corresponding category.

1.1 Background and context

Over the past few decades, the majority of remote sensing studies have been applied to the natural environment. With the development of high resolution imagery, urban remote sensing is gaining interest within the remote sensing community. Among the scientific interests is the development of new techniques of image analysis. The limitations of supervised and unsupervised algorithms in classifying accurately high resolution data have opened up many other techniques using object-based algorithms. Object-based algorithms segment the image into individual objects of unique characteristics based on spectral and shape properties.

However, it has been reported that the single most important technical issue in urban remote sensing is the spatial resolution of the image data (Donnay et al., 2005). Understanding spatial resolution of image data and its influence on object sizes would determine the feature of interest that can be extracted at the appropriate scale (Navulur, 2007).That said, there is no ways of defining a priori what the appropriate scales associated to specific patterns are (Dejong & Van Der Meer, 2005). Still there is a difference between scale and spatial resolution. A resolution expresses the average area dimension that a pixel covers on the ground; scale on the other hand describes the level of aggregation on which a certain phenomenon can be described (Benz et al., 2004). Studying an image at different scale levels instead of a single scale level is a powerful technique to understand the relationships between objects within the image and interpret the scene more easily (Benz et al., 2004; Bock et al., 2005; Taubenbock et al., 2007). A multi-scale approach allows dominant patterns to emerge at their characteristic scales, with no a priori user knowledge in order to obtain adequate and complete information about the patterns (Hay et al., 2003; De Jong & Van Der Meer, 2005).

The sub section that follows investigate the scope and some limitations of the multi-scale analysis approach.

1.2 Scope and limitations

Remote sensing data is said to be dependent on the sampling grid used for their acquisition, and neglecting the object scale and their aggregation level can produce results having little correspondence with the geographical entities within the scene. In fact, there is no unique spatial resolution appropriate for the detection and discrimination of all geographical entities composing a complex natural or man-made scene such as an urban scene. Classification based on unique spatial resolution should be replaced by a multi-scale approach.

The traditional per pixel classification methods with supervised and unsupervised algorithms are said to produce accurate results only when used with low spatial resolution data. With high spatial resolution data it becomes more difficult to extract information accurately, especially in complex environments such as urban areas. Object-based methods which assign group of similar pixels to corresponding classes using objects topology, size and shape information are said to produce more satisfactory results compared to traditional pixel-based methods(Niebergall et. al.,2007; Pu et al., 2009).

Image interpretation is a difficult task which can be defined as the automatic extraction of semantic data from an image.However, this semantic data is not always explicitly available within the image and depend on scene knowledge. Object-based methods involve segmenting images into homogeneous regions which are then described by a set of features related to spectral signatures, geographical features such as shape, length, and topological properties such as adjacency. These features are then used in the recognition and the classification process. Few works focused on the development of such scene knowledge and contextual models for identifying and classifying urban areas.

1.3 Hypothesis

Three hypothesis will be tested in this research:

1. The quality of high level features selected as inputs for classification can predict the classification outputs results.
2. Change in object size results when changing the segmentation scale parameter indicates a relationship between the scale value, the spatial structure of the image and the intrinsic structure of the object generated.
3. The integration of contextual information into a multi-scale classification process is a reliable approach that produces better results compared to single scale pixel-based approach.

1.4 Aim and objectives

The aim of this study is to develop a multi-level context-based classification system applicable to any urban area using high resolution images. To reach this aim, the following are the objectives:

1. Identify, calculate and assess high level contextual and non contextual features in order to build a scene model.
2. Evaluate scene intrinsic structure and object internal variance in order to determine optimal segmentation scale parameters.
3. Develop a classification strategy that integrates scene context in order to accurately assign each image object to its corresponding class.

1.5 Thesis Outline

This thesis is organized as follows:

Chapter 1 presents the problems investigated in this research, as well as methods to resolve those problems.

Chapter 2 which is the literature review,discussed firstly the recent advances in segmentation techniques and secondly investigated some of the mostly used classification techniques. An investigation of feature selection techniques was done at a third stage and finally, a study of the most relevant urban classification studies was done before an overview of some post-classification processing tools.

Chapter 3 will analyse the scene under investigation in order to extract suitable features useful to build a scene contextual model to be used for classification.

Chapter 4 will discuss the segmentation of the image into several hierarchical levels in order to perform a multi-level classification of the study site.

Chapter 5 is dedicated to the multi-level classification of the study site by using object contexts in addition to spectral, size and shape characteristics.

Chapter 6 deals with the conclusion and suggestions for future work.

Chapter 2

Literature review

Understanding a scene, in order to accurately recognize the different spatial objects and their mutual relationships in the real world, has attracted the interest of remote sensing scholars. This interest has led to the development of new image segmentation and classification tools. However, the traditional pixel based classification techniques have demonstrated some limitations when applied to high resolution data. As a consequence, scene knowledge is required in order to improve the classification results by incorporating object descriptive information such as shape, size, context, in addition to its spectral properties. Moreover, several post processing tools are also made available in order to improve the accuracy.

This chapter reviews, in its first section, nine of the most recent segmentation techniques used in remote sensing object-based recognition. Techniques such as Region-Based Level Set, Knowledge-based, Context based, Texture based, Shape and Size-based, Neural Network, Fuzzy Clustering, Temporal and the Multi Resolution segmentation techniques are investigated. In the second section of the chapter, an investigation of some of the most advanced classification techniques is done. Classification approaches such as pixel-based, sub-pixel based, per-field, context-based and knowledge based classification approaches are studied. The third section of the chapter examines some of the most recent scene description techniques available, for instance, techniques to measure object size and shape are studied. The different materials present in an urban scene and their spectral properties are also investigated in this chapter. The fourth section investigates some measures of object context spatial relations such as distance relations. The last section of this chapter will list a few of the post-processing tools that can be used to improve classification accuracy. Techniques such as filtering and manual post classification editing tools are explored in this section.

2.1 Investigation of some advanced image segmentation techniques

Segmentation result can be considered as a key element in object recognition from aerial and satellite imagery. However, this result depends on the segmentation algorithm used, as the algorithms available differ from one another in the sense that they emphasize different desired properties in the object of interest. Segmentation algorithms generate homogeneous and uniform regions with respect to some image properties, such as pixel values, colour, texture, object size and shape. The resulting objects must not overlap and must differ from one another in respect to the properties used to create them . However, several researches have investigated the different image properties in order to implement new segmentation algorithms (Weszka, 1978; Haralick & Shapiro,1985; Spirkovska, 1993). In this section nine of the most advanced segmentation techniques used in object recognition are discussed.

2.1.1 Region-based level set segmentation technique

Region-based level set approach was introduced by Kass et al., (1987). The technique relies on the deformation of initial object contours in order to achieve a desired object delineation. The deformation of the initial contour is operated through the minimization of a function whose minimum is obtained once the desired boundary of the object is achieved. A mathematical description of the technique can be found in Karantzalos and Arglalas (2009). The authors tested the technique on aerial imagery in order to extract man-made features such as road networks, building boundaries and airports. The technique was reported to be efficient in the detection of man-made features under investigation. Road networks, for example, were entirely detected on a 1 meter resolution RGB aerial image. In fact, the approach failed to detect only two roads in the entire network through an over segmentation . Road patterns were detected with an accuracy of 97.32 percent, which is quite good in respect to the resolution of the imagery. In addition, the technique was tested on a panchromatic 0.4 m resolution image to investigate the potential of the approach in extracting building boundaries. While the objects under study were successfully identified, some buildings could not be separated accurately from their shadows. An overall accuracy of building detection was reported at 90 percent. Moreover,when tested on a 0.7 m resolution RGB aerial image and a merge Quickbird image, the technique produced poor detection accuracy due to the influence of shadows, occlusions and spectral variabilities between building roofs more specifically. Finally, when tested on a SPOT High Resolution Visible image, in order to detect an airport, the technique produced promising results as the airport boundaries were accurately extracted and the different objects within the airport scene were discriminated.

2.1.2 Knowledge-based segmentation technique

Segmentation of complex scenes, such as urban scenes, is very challenging without prior object or scene knowledge due to the similarity in respect to spectral and shape characteristics between objects of different categories within the scene. The complexity of certain scenes has led some authors to distinguish two types of segmentations:(1) a so-called partial segmentation, which does not take into account a prior knowledge of the image content and (2) a complete segmentation which takes into account scene characteristics.

Moreover, Ton et al., (1991) tested a two-step knowledge-based segmentation technique on a Landsat image over Fredrick Township in Michigan, USA. The technique performs at a first-stage category-oriented segmentation, using kernel region extraction, and then applies an image-oriented segmentation through a hierarchical segmentation. The two steps rely on spectral and spatial rules.

In addition, the approach was reported to have the potential of dividing the image into homogeneous regions. The degree of discrimination found was 85.7 percent, 86.3 percent and 87.5 percent respectively for deciduous vegetation, coniferous forest and non forest. Moreover, Cicala et al.,(2004) developed a new knowledge-based segmentation technique, that used a tree structure to Markov Random Field, to generate homogeneous regions from aerial images. A mathematical explanation of the method can be found in Poggi et al.,(2005).

2.1.3 Context-based segmentation approach

Context-based segmentation has widely been tested in image recognition (Cheng & Bouman,1998; Cheng & Bouman, 2001; Fan & Xia, 2001). Cheng and Bouman (2001) developed a multi-scale Bayesian segmentation algorithm, that takes into account local and global context within the image. The algorithm developed can be trained by an existing accurate model. The context-based segmentation uses as a first stage, a context model of which parameters are computed from the training image, rather than the image to be segmented. In the second stage, the technique uses a multi-scale image data model and the Haar basis wavelet coefficients as image data structure (Cheng & Bouman, 2001).The correlation between the different wavelets coefficients, through the hierarchical scales, is modelled as an autoregressive process. After the complete model is created, the algorithm uses the sequential maximum a posteriori (SMAP) estimator to partition the image into homogeneous areas. However, the success of the model relies on some segmentation parameters, namely, the context model parameter, the quad tree parameters, the decimation of ground truth segmentation and the estimation of data model parameters. Moreover, transitional probabilities, which parametrize the estimation of the context model rely on a binary encoding function built on the basis of pixels at coarse resolution and a binary encoding function of the unknown pixels. On the other hand, the estimation of the quad tree parameters is computed using the transitional probabilities. In addition, the decimation of ground truth segmentation is based on the quad tree parameters. The estimation of the data model parameters is done using a Gaussian mixture model, to estimate the conditional probability distribution (Cheng & Bouman, 2001). Once all the parameters were computed, the algorithm was tested for document segmentation and the authors reported the technique to be computationally efficient and the segmentation results found were of high quality. Moreover, other segmentation techniques, involving context and multi-scale approach have been investigated and tested on aerial imagery (Fan & Xia, 2001).

The notion of context is generally related to the spatial relationship between neighbouring objects, and most of the context-based segmentation applies Bayesian probability. Interesting documentation, explaining Bayesian theory, can be found in Lacave and Diez (2002). Bayesian segmentation uses Bayesian estimators that minimize the average cost of an erroneous segmentation. One of those estimators, the maximum a posteriori (MAP) estimator, maximizes the probability that every pixels in an image would be correctly grouped into homogeneous and distinct regions. However, the MAP was reported to not be suitable for multi-scale segmentation, due to the fact that the estimator minimizes the cost function, which is an important factor that weights both coarse and fine segmentation levels. Minimizing the cost function at coarse scales can led to an erroneous segmentation, as the level contains many more pixels that can be affected. For instance, 2^{n^2} pixels can be affected at finer scale segmentation if the MAP is used to segment an image at n levels (Cheng & Bouman,2001). As an alternative, to overcome the limitation of MAP, SMAP estimator was used. The SMAP begins by computing coarser level segmentations and at each level, a new segmentation at a level n is computed, based on the previous one at a level $n - 1$

However, most context-based segmentations do not take object texture boundaries into consideration, which can led to heterogeneous objects at texture boundaries. More recently, to overcome this limitation, Fan and Xia (2001) proposed a technique that combines the advantages of multi-scale and multi-context to create image objects. In addition to the SMAP introduced above, the technique uses a context model which is characterized by a context vector, v^n, derived from neighbouring pixels at a coarser segmentation level, n. The authors tested the new technique on remotely sensed data such as aerial photographs, Synthetic Aperture Radar images and the results revealed an improvement in segmentation accuracy through accurate boundary localization and extraction. In contrast, the technique has been reported to be limited in the extraction of objects of very small size such as cars or single trees.

2.1.4 Texture-based segmentation techniques

Texture-based segmentation techniques have been investigated by many researchers (Lucieer et al., 2004; Guo et al., 2005; Sadykhov & Lukashevich, 2008). In image segmentation, colour is an important characteristic to discriminate between the different objects within a scene. However, the use of colour in image classification was reported to have some limitations with high resolution satellite images (Guo et al., 2005). On the other hand, texture measures have been found to be more relevant to discriminate between objects in an image. Texture-based image segmentation techniques use linear transforms and multi-resolution feature extraction (Unser, 1995), fractal dimension (Eiterer et al., 2004), wavelet transforms (Reed & Du Buf, 1993 ; Jain & Karu, 1996) and the Markov Random Field model (Wang & Zhang, 2003). The texture-based segmentation uses a technique called Local Binary Pattern (LBP) which is described in details in Eiterer et al., (2004) and Ojala et al., (1996). The LBP was reported to be robust and suitable to handle multi-resolution images and computationally efficient (Guo et al., 2005). The LBP is created using the joint distribution of values of a circular symmetric set of pixels in a local neighbourhood (Guo et al., 2005). However, the LBP was reported to not be suitable for high resolution imagery as the texture description implemented in the LBP is highly discriminating of texture features and considers the variation of pixel reflectance as a different texture feature.

To investigate this issue, Ojala & Pietikainen (1999) tested a new approach incorporating a feature distribution function. However, the approach was found not to be as suitable for high resolution data as the previous technique. In contrast, Guo et al.,(2005) tested a texture-based approach that filters the image at a first stage, using a wavelet filter, to produce multi-scale wavelet images which are then split at a second stage, based on the quad tree technique, to create multi-level split images. The split images are then optimized to generate the final split image results which are merged to produce the final segmentation results. Clear texture descriptor features used in the technique can be found in Ojala et al, (2002). As high resolution data contains noise, the wavelet transforms used in the approach eliminate image high frequencies, to produce a better resolution of the image and a better frequency, more useful for a successful segmentation.When tested on a high resolution ortho image, the technique was reported to be efficient in respect to the quality of image objects produced and the computational time. However, the approach has some limitations when segmenting the boundaries of objects.

2.1.5 Shape and Size based segmentation techniques

Shape and size are very important characteristics in image classification. In fact, shape can play an important role in discriminating two or more objects of similar spectral responses. For instance, a roof can be distinguished from a road based on their respective shape or size. Road have large $\frac{Length}{Width}$ ratio, characterizing a linear feature, while a building roof has a smaller ratio.

Several authors have used shape to segment aerial images or assess segmentation results (Beveridge et al., 1989; Bongiovami et al., 1993; Pishva et al., 2000). It is possible to improve segmentation results using objects shape information. For instance, Beveridge et al., (1993) improved an over segmentation result by integrating object shape information, using a combination of localized histogram and region merging techniques.

In addition, Pishva et al.,(2000) applied a shape-based segmentation to differentiate various types of bread within a digital photograph. Before segmenting the image, the authors applied a series of pre-processing methods on the image.The methods consisted of a Dark correction, background correction, K highlight correction and colour balancing. After pre-processing techniques, the data was binarized in order to extract shape and size information, such as area, elongation and minimum bounding box. After extracting shape information, a texture analysis was performed using the dependency matrix technique. Moreover, the technique performed a colour analysis in two consecutive steps. The first step involves the histogram analysis of the hue component of the bread sample and then, at a second stage, the histogram analysis is applied on segments of similar colour to segregate similarly coloured breads from other segments of different colour distribution (Pishva et al.,2000). To complete the segmentation process, the hue histograms were subdivided into fourteen colour groups, and areas of each colour group was computed. The area values found were then used to create the distribution matrix in order to auto-scale the image and calculate a unity standard deviation (Pishva et al., 2000). The technique produced promising results as all the bread types were successfully segmented. The technique can be used with any orientations and can detect circular as well as elliptically shaped objects.

2.1.6 Neural network segmentation techniques

Most of segmentation approaches used today are supervised techniques, but have been reported to have some limitations in the sense that they require prior knowledge about the data, in order to perform successfully (Awad, 2010). To overcome the limitations, several unsupervised segmentation techniques were tested, and among them are the neural network techniques. The approaches have been widely investigated in image segmentation (Beneditksson et al., 1990; Zhang et al., 2003; Awad, 2010). Among the neural network techniques mostly used is the Self Organization Map approach (SOM). The SOM has been investigated by many researchers (Yin & Allison, 1995; Kohonen, 2001; Zhou et al., 2007). Aria et al.,(2004) have previously tested the technique on satellite imagery. Awad et al.,(2007) combined the SOM with an hybrid genetic algorithm to segment various satellite images, but the technique was reported to have limitations in computational time. Moreover, Zhou et al.,(2007) tested another algorithm combination segmentation, involving the SOM and the K-mean algorithm. This combination technique had the advantages of segmenting images at coarser and finer scales respectively using K-mean and SOM. But the techniques shown some limitations as the K-mean algorithm requires a pre-defined number of classes to be used in the segmentation. Another limitation of the technique is the long computational time.

More recently, Awad (2010) proposed a method, that combines the SOM and a new algorithm called the T-cluster.The SOM technique converts arbitrary dimensionality of patterns into a two dimensional array of neurons

(Awad,2010). Neurons in this technique constitute an orthogonal grid of cluster units which are each linked to three internal weights applied to the three layers in the case of a three channel aerial image.The approach uses the minimum Euclidean distance, to select the best cluster unit that matches the input pattern. The selected unit and a certain number of neighbouring pixels are taken into account for the calculation of new internal weights in order to find the closest value to the input model. This technique was reported to have limitations in segmenting small regions, which have a few number of pixel candidates. However, this limitation can be improved with the use of T-cluster in the process. The T-cluster has the capability of eliminating small clusters and limit over segmentation occurrence during the segmentation.The selection of best clusters by the T-cluster is done in three steps which follow :

1. The clustering process starts by calculating the distance between the values of the cluster centres.

2. Two clusters are merged, if the distance between them is smaller than the pre defined threshold T.

3. If the step (2) is satisfied, then the approach considers the minimum number of pixels in the merge process. This means that, clusters with smaller number of pixel are merge to the larger ones.

Moreover, SOM and T-cluster algorithm complete the segmentation in a sequential way. SOM organize pixels in clusters, so that the highest peaks of the histogram are considered as cluster centres that are then captured by the T-cluster, to generate the final segmentation results. The performance of this technique was measured against the Iterative Self Organizing Data (ISODATA) algorithm. The approach was reported more efficient than the ISODATA, in respect to the quality of segments produced and the computational time.

2.1.7 The Fuzzy clustering segmentation techniques

The main goal of most segmentation techniques is pixel clustering. Clustering techniques have been widely investigated by researchers (Jain & Dubes,1988; Everitt,1993; Pal & Bezdeck,1995; Maulik & Bandyay,2002; Bandyopadhyay,2005). Clustering techniques divide the image into K regions, based on certain similarities. Several clustering algorithms are made available in literatures and among them are the Fuzzy C-means, the Average Linkage clustering and the Simulated Annealing based Fuzzy clustering.

The Fuzzy C-means uses the theory of fuzzy sets to segment the image. The algorithm starts with randomly selecting K centres and then, at every iterations,determines the Fuzzy membership of each pixel, until there is no change in the cluster centres. In addition, once the cluster centres are stabilized, each pixel is assigned to the group with which it has the highest membership value. A detailed description of other clustering algorithms such as Simulated Annealing and Average Linkage clustering can be found in Maulik and Saha (2009).

Another fuzzy clustering algorithm, the Differential Evolution, has widely been tested for for image segmentation (Storn & Price,1995; Price,1996; Price & Storn,1996; Storn & Price 1997; Price & Storn,2005). More recently, Maulik and Saha (2009), tested a new Differential Evolution technique on satellite imagery. The approach optimizes the image in continuous domains, and the decision variable is represented by a real number. Difference Evolution algorithms randomly generate initial clusters, and then, evaluate them. After the evaluation, a process called mutation takes place, in order to create a simple offspring which competes with a parent, to determine the most qualified parent to pass to the following generation (Maulik & Saha,2009).

Furthermore, the single offspring is generated by addition of weighted difference vector between two parent vectors to a third parent vector. This step is called the crossover in which, each offspring and parent vectors are used to create trial vectors, which depend on the crossover rate that can be defined by the user in a range between 0 and 1. In addition, if the trial vector generates a lower objective function value than a predefined cluster number, the newly generated vector replaces the vector with which the comparison was operated (Maulik & Saha, 2009). A detailed description of the Differential Evolution based Fuzzy clustering is developed in Maulik and Saha (2009). The new technique was reported to be a success when tested on satellite imagery of Calcutta. The method enabled an accurate segmentation of the different predominant classes within the scene. This success originates from the optimization of cluster validity measure. However, the technique was revealed to have some drawbacks in automatically generating suitable number of clusters, to be considered in the segmentation process. Other authors have previously tested some variations of the method (Zhang,2006; Polenok & Sadykhov,2008). More recently, Sadykhov (2008) tested a combination of Fuzzy clustering segmentation with a non linear filtering algorithm on Landsat multi spectral images. The technique accurately allowed discrimination between the different land covers under investigation such as wetland, water-meadow and bush areas. In contrast, the method has limitations in defining the number of initial clusters, as well as the appropriate cluster centres, which rely on the number of iterations.

2.1.8 Temporal image segmentation techniques

Most of temporal segmentation techniques have been used in the detection of changes in natural and man made environments (Jeansoulin et al.,1981; Carlotto,1997). Initial investigations on change detection were introduced by Resenfeld (1961). The investigations were based on statistical similarity measures. These works stimulated further developments of segmentation techniques in order to detect changes. Among those works is the research done by Price and Reddy (1997). Other temporal segmentations techniques such as adaptive subtraction, linear prediction and perpendicular change index have been respectively explored by Lee et al.,(1986), Therien et al., (1986) and Carlotto et al.,(1997). Temporal segmentation partitions images into unique regions of change across the image over time (Carlotto,1997). The technique thresholds the total difference images, in order to generate a binary change image. Patterns of change are represented by an image of labelled vectors. In addition, two labelled vector are considered equal, if there is no change in respect to a pixel characteristics between the corresponding times. In cases where changes have been identified, temporal filtering is used to enhance and detect the changes. Temporal filtering associates temporal images, such that those patterns of change are emphasized. The filter adds the absolute difference that occurs when changes are expected and then,subtracts the absolute differences that occur when no changes are expected (Carlotto,1997). After a test of the technique on a Landsat image, promising results were found. The method overcomes the limitation of the linear model-based segmentation techniques, when large changes occur within the image. Applied on all Landsat bands, the method proved its potential to discriminate areas of no change from areas of changes. Other temporal segmentation techniques, using Fuzzy edge detection and Fuzzy region growing, have been developed and details can be found in Jeansoulin (1981). Another approach using temporal segmentation, involving spatial knowledge, can be found in Hanazumi et al., (1991).

2.1.9 Multi resolution segmentation techniques

The importance of scale in the identification of objects within satellite images was firstly introduced by Woodcok and Strahler (1987). Scale can be defined as a level of aggregation at which an object can accurately be described (Benz et al.,2004). Some researchers pointed out the difficulty of accurately extracting at a single level the diversity of objects of various size. As a consequence, several researches on a multi-scale approach were undertaken. However, the multi-scale object extraction goes with multi -resolution segmentation techniques. Multi-resolution segmentation techniques have been widely used in image classification (Bouman & Shapiro, 1994; Gross et al.,1994; Fosgate et al.,1997; Chen et al., 1997; Krishnamachain & Chellappa, 1997; Moulin & Liu, 1998; Nowak,1999; Li et al., 2000; Choi & Baraniuk, 2000; Venkatachalan et al.,2000). Among multi-scale segmentation techniques widely used, is the Fractal Net Evolution technique, which was introduced by Baatz and Schape (2000). The Fractal Net Evolution technique considers each pixel as one image object. The technique merges neighbouring objects to the seed object pair wisely, based on a certain number of homogeneity criteria. In addition to the homogeneity criteria,a merging cost is assigned to different merge processes (Baatz & Schape,2000). No merge is possible unless the merging cost or degree of fitting is evaluated, and found smaller than a certain least given cost.When there is no more possible merges, the segmentation stops. A larger degree of fitting enables a large number of merges and, smaller degree enables fewer merges. Moreover, the size of resulting image objects depends on this degree of fitting. Because of this property, that segmentation parameter was named scale parameter(Baatz & Schape,2000). The Fractal Net Evolution segmentation has two components which are, the decision heuristics, which determines the image objects that will be merged at each segmentation step, and the definition of an homogeneity of image objects through the computation of the degree of fitting between two neighbouring image objects. Several possibilities for decision heuristics can be used, for the choice of the best object which will take part to the merge:

1. The fitting possibility that merges an object A with any neighbouring object B that fulfil the homogeneity criteria.

2. The best fitting that merges an object A with its neighbour B which fulfil the best homogeneity criteria.

3. Local mutual best fitting enables the object A to find its neighbouring object B which fulfil the best homogeneity criteria and confirms that the homogeneity criteria is best mutually fulfilled.

4. The Global mutual best fitting that merges a pair of image objects fulfilling best the homogeneity criteria within the whole scene using a distributed treatment order of image objects.

The Fractal Net Evolution considers a pixel as a fractal and, represents the concept of hierarchy as a fractal net. The technique has the potential of providing fractal spectral, texture, spatial, shape, size and context information. Tested on satellite imagery ,the method was reported to produce highly homogeneous image objects (Baatz & Schape,2000). In contrast, the technique revealed limitations in accurately segmenting objects edges. A complete mathematical description of the method can be accessed in Choi and Richard (2001), and Zhang et al., (2005).

Similarly to other segmentation techniques, multi-resolution segmentation also has limitations, especially in defining suitable segmentation parameters. As a consequence, several authors have investigated possibilities of defining appropriate segmentation parameters. For instance, Maxwell and zhang (2005) proposed a technique which automatically determines segmentation parameters of multi-resolution segmentations. Kosir and Tasic (2002) proposed a pyramid segmentation parameters estimation technique based on total variation of the image. Moreover, Salvador and Chan (2004) proposed a technique for determining the number of clusters in a multi-scale segmentation. Kim et al.,(2008) on other hand proposed some improvements to the technique introduced by Woodcok and Strahler (1987). Instead of computing the standard deviation of the image from a 3×3 moving window, they derived it from objects obtained through segmentation. However the technique was reported to be suitable to define only a single optimal scale.

To introduce a multi-scale dimension into the technique made available by Woodcock and Strahler(1987), Dragut et al., (2010) proposed a tool called the Estimation of Scale Parameter (ESP). Tested on Light detection and ranging imagery, colour photography and Quickbird images of respectively mixed residential and forested areas, natural savannah and a temporary settlement area, the technique produced promising results. For mixed residential and forested areas segmented at scales of 14, 45 and 82, the technique correctly differentiated objects within their respective classes. The test done on savannah produced a good separation between trees and shrubs at a scale of 16 . The forest stand was captured at scale of 36 and discrimination between forest, grass was possible. Discrimination between bare soil and shrubs was reached at scale of 88. Moreover, temporary settlements were segmented with scale parameters ranging from 1 to 50. The scale parameter of 18 enabled identification of dark huts and scale parameter of 35 enabled accurate segmentation of tents. Even though, the technique produced high quality segments, the results were reported to be influenced by shape and compactness parameters, and further researches combining scale parameters, smoothness, colour and shape compactness should be investigated in order to determine the appropriate parameter combination.

2.2 An investigation of some advanced classification techniques

Since the availability of high resolution aerial images, several advanced classification techniques such spectral mixture analysis, Fuzzy set classification, Artificial Neural Network, Maximum Likelihood, ECHO classifier, and object oriented approach have been applied to image classification. In this section, some of the most used advanced techniques were grouped into five categories in respect to the domain they operate in order to classify images. The pixel-based, sub-pixel based, per field, contextual and knowledge-based classification techniques will be investigated.

2.2.1 Pixel-based classification techniques

Pixel-based classification methods use pixel signatures to assign each individual pixels to their corresponding categories. Most of pixel-based techniques use class samples to classify scenes, but class samples should be representative enough, in order to take into account all the spectral variability within the class. Pixel-based classifiers can be grouped into two subcategories : the parametric classifiers and the non parametric classifiers (Lu and Weng,2007). The parametric classifiers require some statistical parameters such as mean vector or covariance matrix, computed from the training samples. Parametric classifiers require representative samples for each category involved in the classification process. As example of mostly used parametric classifiers are the Maximum Likelihood, the Linear Discriminant Analysis and the Minimum Distance. In contrast, non parametric classifiers do not require any statistical parameter to assign pixels to corresponding categories and their advantage of integrating non spectral data can improve classification accuracy. The Neural Network technique, Support Vector Machine, Decision Tree Classifiers,and Evidential Reasoning, are among the mostly used non parametric classifiers. The Neural Network approach has widely been used for image classification (Beneditksson et al.,1990 ; Foody,1997). A detailed description of the technique can be found in Foody (1997).

2.2.2 Sub-pixel classification techniques

Techniques based on pixel values have been reported to have limitations in accurately separating different objects within a scene due to spectral similarities. To overcome that disadvantage, sub-pixel techniques such as Imagine Sub-Pixel Classifiers, Fuzzy Expert Systems, Fuzzy Neural Network and Linear Regression, have been proposed. Several authors have tested different Fuzzy approaches to classify aerial images(Bellman & Zadeh,1987; Chameau & Santamaria,1987; Binaghi,1992; Rocha & Yager,1992; Binaghi,1993; Binaghi et al.,1993; Binaghi & Rampini,1993; Binaghi & Montesano,1994; Binaghi et al.,1997; Chang & Cho,2002; Salman et al.,2008). Fuzzy classification techniques use membership functions, to assign objects to their appropriate categories. However,two of the challenges in the use of membership function is the definition of class boundaries, and the selection of appropriate objects features to be used in the classification.

Furthermore, Fuzzy techniques can be combined with other algorithms, in order to improve classifier performance. Salman et al.,(2008) tested two different Fuzzy approaches, one based on membership functions, and the other based on nearest neighbour algorithm. Both techniques were tested on images of different resolution. The 30 m resolution Landsat data, the 10 m resolution SPOT 4 data and the 4 m resolution Ikonos were considered in the investigation. The classification results revealed that the Fuzzy membership function produced higher classification accuracy than the Fuzzy nearest neighbour technique, when tested on Landsat image. In contrast, Fuzzy nearest neighbour produced higher results when applied to SPOT4 high resolution visible image. Similar results were found with Ikonos image. However, the study revealed some limitations from both techniques in discriminating buildings from tree shadows. A similar test comparing Fuzzy membership approach and Fuzzy nearest neighbour was done by Chang and Cho (2002).

The consideration of contextual information in Fuzzy classification have also been studied. Binaghi et al., (1997) tested a new approach combining fuzzy logic and contextual information, in order to classify three land cover classes namely snow, ice and others. The technique produced different images of the different classes involved in the classification by partitioning the spectral space of the image (Binaghi et al.,1997). In the spectral space, each pixel value represent a degree of membership in the corresponding class and the membership value are proportional to the degree to which the targeted pixel contains the corresponding land cover (Binaghi et al.,1997). To add contextual information within the classification process, the research used a tool called the Fuzzy PI Granules, which defines object context. Tested on a Landsat image, the technique was reported to have produced good results. For instance, an overall classification accuracy of 88 percent was found for the Orthes Ceveledale class, and 87 percent for the South Pusteresi Alps (Binaghi et al,1997). The fuzzy logic technique used in this study was reported appropriate for object extraction, due to its flexibility, which minimizes the effect of mixed pixels at class boundaries by capturing the intrinsic and spectral vagueness characterizing those boundaries.

2.2.3 Per-field classification approaches

In complex scenes such as urban environments, the spectral variability within roof group can lead to spectral confusion with other classes such as roads or bare soil. Classifying such environments using pixel or sub-pixel techniques can produce inaccurate classification results. To overcome this limitation, per-field classification techniques have been investigated (Aplin et al,1999 a,b; Alplin & Akinson,2001; Dean & Smith,2003; LIoyd et al,2004; Lewinski & Bochenek,2008). The approaches use land parcels instead of individual pixels. The land parcels used in the technique can be produced using GIS tools and integrated into the segmentation. However, the consideration of vector layers to train the segmentations, can be an obstacle, as they are not always available, and are difficult to produce and update for changing environments such as urban environments. This limitation can be overcome with the use of new per-field classification techniques such as the Object Oriented approach. The technique is composed of two steps, the segmentation and the classification. As introduced in the previous section, segmentation groups similar pixels, to create homogeneous units called segments. Those homogeneous regions do not only contain spectral information but also shape , size, texture and context. Moreover, the quality of segments generated rely on parameters such as scale parameter, colour, shape compactness and smoothness.

The notion of image objects has been initially introduced by Gonzalez and Wintz (1997), and from that period new researches were undertaken in order to develop techniques and tools for object-based classification. Object-based classification has been widely used for remote sensing applications such as detailed and broad mapping projects of man-made and natural environments (Herold et al., 2002; Asmat et al., 2003; Herold et al., 2003; Marangaz et al., 2004; Misakova et al., 2006; Yan et al., 2006; Jianyu et al., 2006; Niebergall et al., 2007; Yue et al., 2008; Li et al., 2008; Chen et al., 2009; Huang and Qi , 2009; Blaschke, 2010; Stow et al., 2010).

Object-oriented approach considers the notion of hierarchy levels in such a way that upper levels are created from lower levels, creating objects topology. A mathematical description of the concept, as well as the different parameters involved, can be found in Jianyu et al., (2005) and Li et al., (2008). Niebergall et al., (2007) tested the object-oriented technique on an urban scene of Delhi in India, using a merged Quickbird image. The test revealed the potential of the approach to detect the different settlements types compared to pixel-based statistical techniques such as Maximum Likelihood. Huang and Qi (2009) tested the potential of the object-based method in mapping the natural reserve of Changbai in China, using an Ikonos image. The measured classification accuracies were found greater than 81 percent for the different land covers namely tundra, bare rock, mountain birch and volcanic lake.

A comparative study applying object-oriented technique on high resolution imagery was done by Herold et al., (2003). The authors tested the classification approach on an Ikonos image of Santa Barbara. The fuzzy membership method used was based on fuzzy class descriptors, involving spectral and spatial object features. The classification was performed using three levels to extract water, built up, vegetation and non vegetation. Spectral information could not be enough to discriminate the different classes. As a consequence the authors investigated the contribution of size and shape information to the process, and the results found were very promising, as bare soil could be separated from beach and bare rocks, with which they share similar spectral properties. After classifying the main classes at level 1, roofs were separated from transportation infrastructures, based on spectral and shape information at level 2. The compactness ratio was for example used to discriminate roof from other features such as roads as roofs are more compact features than roads. In addition, spectral information was supportive, to extract individual objects at level 3 classification.

The approach was reported to be a success in mapping complex environments such as urban environments and an overall classification of 79 percent was found. Another application of object-based classification method is the detection of land cover heat over an urban area (Asmat et al.,2003). The research used an enhanced Landsat TM image, and applied a multi-resolution segmentation on the data. Surface temperatures of different land covers were estimated using the Normalized Difference Vegetation Index (NDVI).

Once temperature was estimated, a fuzzy membership classification was used to classify different land covers according to their respective temperatures. The technique produced interesting results characterizing different land cover temperatures. For instance the technique enabled measurement of a temperature of 25 degree Celsius for the forest land cover. Agriculture fields were characterized by a temperature of approximately 30 degree Celsius but built up areas were found with the highest surface temperature. Additional per-field classification techniques such as Map-guided classification, Graph-based technique, Structural Pattern Recognition System and Spectral Shape Classification can respectively be found in Chalifoux et al., (1998), Bernsley and Barr (1997).

2.2.4 Context-based classification techniques

To overcome the limitations of traditional pixel and sub-pixel based classification techniques, new classification tools have been tested, and among them are the contextual classifiers(Gong & Howarth , 1992; Kartikeyan et al., 1994; Flygare, 1997; Hinz & Baumgartner, 2000; Cheng & Bouman,2001; Lira & Maletti, 2002; Keuchel et al., 2003; Singhal et al., 2003; Magneussen et al., 2004; Porway et al., 2008; Liu et al., 2008).

Objects context in an aerial image refers to the spatial relationships between the different objects within the scene. However, contextual information can be categorized into three groups, namely, topological relations,distance relations, and relations of cardinal direction. A description of the different objects contextual relationships can be found in Liu et al.,(2008). Contextual classification techniques use context models and a review of some recent techniques useful to build contextual models can be found in Battle et al.,(2000). The authors provided models for the recognition of objects such as houses, roads within a scene. But the proposed models had limitations in the sense that they can only be applied to a single type of object (Singhal et al.,2003). Singhal et al.,(2003) proposed a new tool called probabilistic spatial context model that uses the different spatial relationships between objects within an image. The approach defines spatial relationships between objects based on two ways: the first way uses the bounding box of the region under investigation, and the second technique relies on a lookup table of directional weights of the region under study. The directional weights of regions are calculated based on the method introduced by Barretti et al.,(2002). Applied on a colour image, the technique produced interesting results as the addition of contextual information into the classification improved the accuracy from 70.6 to 95.2 percent. More context-based classification techniques are available in literatures and among them are: the Extraction and Classification of Homogeneous Objects (ECHO) (Landgrebe, 2003; Lu et al.,2004), Point to Point Contextual Correction method (Cortijo & de la Blanca,1998), Fuzzy Contextual Classifier (Binaghi et al.,1997), Frequency-based Contextual Classifier (Gong & Howard,1992, Xu et al.,2003), the Hierarchical Context-based Classifier (Porway et al,2008), Bayesian-based Contextual Classifier (Cheng and Bouman, 2001).

Moreover, Lira and Maletti (2002) tested a new supervised contextual classifier based on region growing technique, to classify SPOT images. The new approach requires training data in order to assign pixels to their corresponding classes. The approach begins with the identification of seed pixels representing each class and then, remaining pixels are merged to the closest seed sample with which they satisfy the homogeneity criteria the best. In addition, once the training data is generated, the technique employs a pixel centred window to compute the density function characterizing a pixel type in the image (Lira & Maletti, 2002). The computed density function is then compared to the density function characterizing the different training samples, and from the comparison, pixels are assigned to their corresponding classes, with which the similarity is the best. However, pixel centred windows are reported to have some limitations in the sense that pixel contained within the window might not always be homogeneous and this generally happen at class boundaries. But the new technique, takes into account those limitation into a separate processing step. Once identified , pixels at class boundaries are reclassified to the rest of pixels, to form the final classification results (Lira & Maletti, 2002). The results found revealed that the classification performed well, compared to pixel-based techniques. To evaluate the classification performance, the authors used the k-coefficient technique proposed by Landis and Koch (1977) and Smits et al.,(1999). The K-coefficient of the classified SPOT image was reported 0.98823 which is a very good results. Even though the technique produced good outcomes, the approach relies mostly on the choice of seed pixels, and the accuracy of the pixel centred window, which is very subjective.

2.2.5 Knowledge-based classification techniques

To deal with the limitation of pixel-based methods, knowledge-based classification techniques have been investigated and techniques such as Rule-based Syntactical approach, visual Fuzzy Classification ,based on explanatory and iterative visualization techniques, multi Temporal classification relying on decision fusion, Supervised Classification based on ongoing learning, are available through literatures (Lu and Weng, 2007). Knowledge-based classification techniques use additional data such as land cover maps, elevation data, population data, road networks, precipitation and temperature data to improve classification results (Lu and Weng, 2007). For instance, building and road density were reported to be useful by the authors to distinguish between commercial and high density residential areas. Ceccato et al., (2005) used rainfall and temperature data to classify malaria high risk areas in Sub-Saharan Africa region, and the integration of such knowledge into the classification improves classification accuracy as malaria vectors are likely to develop under certain temperatures and water bodies. For instance, high presence of the plasmodium falciparum have been accurately classified based on warmer temperatures. The presence of plasmodium vivax could be monitored using slightly warmer temperatures.

Other knowledge-based classification techniques use high level image and object information organized into a system called a semantic network. A semantic network defines a set of concepts such as swimming pool, parking plots, roofs as well as their characteristics. Buckner et al., (2000) tested a knowledge-based classification technique, which integrated existing scene knowledge. The technique called GeoAIDA stored a priori scene knowledge in a semantic network. The different nodes of the semantic model were organized in an hierarchical manner such that each node had exactly one superior node, and the top most node as the scene node (Buckner et al., 2000). Each node in the network could be assigned attributes such as category name and associated with top down and bottom up operators. Top down operators are described by the authors as external image processing operators, that segment the input image, and assign the generated objects to their corresponding classes. The external image processing operators describe the different image objects using a binary mask in order to create a list of labelled regions (Buckner et al, 2000). An example of external processing operators used in the technique is the variance, which enables the distinction of different types of objects within the scene. On the other hand, the bottom up operators organize objects generated by the top down process into a small number of super objects groups and generates a list of hypothesis nodes describing geometric location of objects within the scene. From the external operator processing,is generated a semantic network describing the different objects properties such as size,topology, which can be integrated into the classification process.

Tested on a laser scanned image and an ortho photo, the technique was reported to have advantages over pixel-based approaches, by successfully discriminating for example inhabited areas from agricultural land. However, the approach was reported to have weaknesses in the use of spectral a priori knowledge as agricultural areas have been confused with acreage, meadow and forest areas (Buckner et al.,2000). A similar investigation using pre defined scene and object knowledge can be found in Kunz et al., (1998). Marinov and Zheliazkova (2005) proposed an interactive tool for knowledge-based classification using a priori semantic network and details on the algorithm can be found in the mentioned literature. Several semantic networks exist and can be grouped into six groups describing their respective functions. The definition semantic network describe the node subtypes using relationships such as IS A ,between the nodes. These types of semantic networks are called generalization hierarchy, which preserve the main node properties(Marinov & Zheliazkova, 2005). The Assertion Networks are used to declare propositions. In these type of networks, information is supposed to be contingently true (Marinov & Zheliazkova, 2005). Implication semantic networks on the other hand use implication as the primary relation in order to connect the different nodes within the network. These types of network are generally used to represent pattern of casualty. Moreover, executable semantic networks include some mechanisms such as marker passing or attached procedures which can pass messages or search for pattern and associations (Marinov & Zheliazkova, 2005). Those types of networks process mechanisms that can cause some changes within the network. The Learning semantic network types build and extend the network representation by acquiring knowledge . The last group of semantic networks uses combination of two or more of the previous techniques and more details related to these knowledge representation can be found in Marinov and Zheliazkova (2005).

2.3 Investigation of some of the most recent scene description techniques

Scene description involves the definition of scene knowledge in order to be used in applications such as image classification. The concept involves the calculation of indices that describe spatial objects, their spatial distribution, as well as their radiometric characteristics. To describe a scene several techniques have been tested to estimate objects size,shape, as well as radiometric attributes. The subsections that follow, will investigate some of these techniques, used to describe objects.

2.3.1 Size measurement techniques

The measurement of object or pattern size have attracted the interest of scholars in the early sixties (Dacey,1965).Size measurement have been used in several applications involving the detection of objects and patterns(Kumar et al,2008). Object size is one of the basic information needed for object recognition in complex environments such as urban environments, using advanced techniques such as knowledge based classification approach. The size measure can be considered as an important parameter to quantify and describe a spatial feature. For instance a block of flats and a single house building are both part of human settlement class but both objects can be differentiated based on their respective size as single house building is characterized by a small size and a block of flat is generally characterized by a larger size. One of the most used size characteristics is the polygon area measure, and several techniques are available to estimate a polygon spatial area. The most accurate but difficult and time consuming technique used is the manual extraction through digitizing (Kumar et al.,2008). As an alternative to the manual digitizing ,the point sampling has also been used, but the approach was reported to be less accurate. Size attribute can also be used to discriminate between patterns. For instance, a pattern composed of small size buildings may not represent the same geographical pattern as a pattern composed of buildings of more important size. More recently, William and Wentz (2008) developed a new technique called TOSS, which used objects size attribute to classify them into homogeneous groups within a given pattern. The results was a combination of overlaid layers representing the respective identified group of objects. Tested on four types of data, namely a Quilt colour pattern, a Pennsylvania geological pattern and two land use patterns, the results revealed the identification of similar objects within the investigated patterns. By using the size standard deviation, the approach could determine similarities between objects within the pattern. For instance, a low size standard deviation value described similarity between polygons within patterns. On the other hand, a high size standard deviation described dissimilarities between objects composing a pattern.

2.3.2 Shape description techniques

Objects on the earth surface can be analysed based on a certain number of properties of which, one of the most relevant is the shape (Lee & Salle, 1970). However, shape cannot be expressed as a number like other polygon measurements such as area or perimeter but the terms such as elongated, L shaped, nearly rectangular, roughly circular should be replaced by more accurate index in which each shape would be attributed a number. But this attributing function should satisfy three properties which follows : (1) each shape should be assigned a unique number,(2) two shapes should not be assigned the same number, (3) two similar shapes should be assigned numbers that are close to each other.

Shape is a fundamental spatial characteristics describing the spatial object (Maceachren,1985). One of the most used shape property is the shape compactness measure, which was introduced for the first time in 1822 by Carl Ritter (Maceachren,1985).Following this, several techniques have been used to estimate objects shape compactness.Among those techniques are the perimeter to area ratio, the parameters of related circles,the direct comparison to standard shape and the dispersion of elements of area.

The perimeter to area ratio technique

The perimeter to area ratio is reported to be a suitable technique for simple geometric shapes such as circles, rectangles, squares or triangles, but the techniques have limitations in the sense that the ratios vary with size. To overcome that weakness, researchers suggested the use of square root of the area or the perimeter (Maceachren, 1985). Squaring the area factor in the ratio produces a compactness value ranging from 0 to 1. In addition,the most suggested ratio is in the form of :

$$Compactness = \frac{\sqrt{Area}}{0.282 \times perimeter} \tag{2.1}$$

More details on the calculation of the constant 0.282 can be found in Maceachren,(1985).The equation above gives values closer to zero for circular shapes and closer to 1 for less compact shapes. Moreover,some researchers have modified the previous equation by squaring the numerator and the denominator to produce the following equation :

$$Compactness = \frac{Area}{(0.282 \times perimeter)^2} \tag{2.2}$$

.

In addition to perimeter to area ratio techniques, the parameters of related circles approach has also been widely used to characterize objects shape and the following subsection will review some of the applications of this technique.

Parameters of related circles technique

The parameters of related circles technique was introduced for the first time in 1892 by Ehrenberg (Maceachren,1985). It measures object shape compactness by comparing its spatial area to the smallest circumscribing circle.The area of circumscribing circle can be calculated from the polygon longest axis or from the diameter using the equation that follows :

$$Compactness = Pi \times (0.5 \times longest\ axis)^2 \tag{2.3}$$

The equation above produces values ranging from 0 to 1 with values closer to 1 characterizing more compact shapes. A variation of the technique is the equation that follows:

$$Compactness = \frac{Area}{Pi \times (0.5 \times largest\ area)^2} \tag{2.4}$$

Moreover, another shape index was proposed by Schumm (1963) but the technique was reported to be only suitable to characterize shape of drainage basins.The index is defined with the equation below :

$$Compactness = \frac{2 \times \sqrt{\frac{Area}{Pi}}}{Longest\ axis} \tag{2.5}$$

A variation of the technique is given by the following equation :

$$Compactness = \frac{4 \times Pi \times Area}{(Perimeter)^2} \tag{2.6}$$

The direct comparison to standard shape

Shape compactness can also be determined by measuring the difference between a unknown shape and a standard model. This technique was introduced for the first time by Lee and Salle (1970). The potential of the technique in estimating the squareness, circularity or rectangularity of shapes was tested by the authors on 25 Sudanese villages located along the Nile river. The authors drew village outlines around each village and then compared them to four standard models namely square, rectangle, triangle and circle. A more detail mathematical description of the technique can be found in Lee and Salle (1970). The results revealed respective mean circularity, squareness, triangularity and rectangularity values of 0.511, 0.527, 0.501 and 0.428. The mean shape index for rectangle was found the lowest, indicating a closest fit between village outlines and the rectangular model. Sixty percent of the villages were found nearly rectangular. Twenty percent of the remaining villages were found nearly circular with a low circularity index and the last remaining villages were close to triangle shape. No village was found with a shape close to the square model.But,the accuracy of this technique relies on the accuracy of the object outlines drawn, which can create regions of mismatch between the models and the outlines. The following subsection will study some alternative techniques to overcome this limitation.

The dispersion of elements of area.

Shape measures investigated previously consider shapes as a whole but do not focus on a single parameter of that shape (Lee & Salle, 1970). For instance, the area of mismatch between the shape under investigation and the model is not taken into account within the index calculation procedure. Two measures have been proposed as alternatives to overcome the limitation. Each measure consider a shape as a series of infinitesimally small elements of the polygon area dA. The dispersion of these elements around a shape centroid such as the variance, standard deviation or moment is the basis of such approach.

Related to the shape moment ,the index can be described by the following equation :

$$Dispersion = \frac{Area^2}{2 \times Pi \times \int r^2 dA} \tag{2.7}$$

The equation can also be written as follows :

$$Dispersion = \frac{Area^2}{2 \times Pi \times \int r^2 dxdy} \tag{2.8}$$

with A as the polygon area and r the distance from the shape centroid to a particular element of the area dA. The terms dx and dy represent the length and the breadth of the infinitesimally small rectangles. A derived equation from the above is given as follows :

$$Compactness = \frac{Area}{2 \times \Pi \times (Q_x^2 + Q_y^2)} \tag{2.9}$$

More literature on the disadvantage of these measures can be found in Massam and Goodchild (1971).

2.3.3 Investigation of urban materials and their spectral properties

Urban materials

Many researches have investigated urban environment in order to determine the different urban materials (Weng & Quattrochi, 2007; Herold et al., 2004; Herold et al., 2002; Robert et al., 2003). But only a few systematic researches focussed on the spectral properties of urban materials which can vary from an informal to a formal settlement area (Busgeeth et al., 2008). To study urban materials , Weng & Quattrochi (2007) investigated the Santa Barbara urban environment and found out that it is composed of rocks, soils, green vegetation, non photosynthetic vegetation, as well as artificial surfaces such as concrete, asphalts, paints, plastics, and metal surfaces. At a more detailed level the study found that some urban land cover components have similar material properties and consequently could produce a similar spectral response. They reported that shingle or tar roofs and asphalt roads, for example, that are different land cover types, produced the same spectral response.

The degree of detail of urban features on an aerial image is strongly related to the spatial resolution of the sensor (Herold et al., 2002). To study the relationship between urban materials and spatial resolution, Herold et al., (2002) classified urban environment into three different hierarchical levels. The level I representing main land cover types like vegetation, built-up, or artificial surfaces and water bodies. The level II subdivided the level I classes based on their use and function. For built-up materials as example, this level discriminates building roofs and transportation areas from others. The level III represented a further stage of class detail based on materials properties for built-up classes. This level shows different roofs and road materials and colours. The study used roofing materials as a function to discriminate building types. It was also found that parameters such as atmosphere, vegetation fouling, object structure, and geometry can influence the spectral response of urban materials. They reported, for instance, that a part of roofs might be illuminated while others remain in shadows due to the geometry or the orientation of the roofs with regard to the position of the sun.

Other researches have even investigated the chemical and physical characteristics of urban materials (Goetz et al., 1985; Hepner et al., 1998; Hepner & Chang 2001; Ben-Dor et al., 2001). The study done by Ben-Dor et al., (2001) reported that the physical and chemical characteristics of different urban materials are represented in all parts of the visible, near infra-red, short wave infra-red, and thermal infra-red spectrum. In contrast, other studies revealed that the spectral discrimination of urban materials is especially difficult at coarse spectral resolution (Sadler et al., 1991; Herold et al., 2004).

Although these studies provided important insights about urban materials, a comprehensive study of the quantitative assessment of their spectral separability and an evaluation of which wavelengths are most suitable for spectral separation is still lacking (Harold et al., 2004).

Urban spectral properties

Urban spectral complexity has been studied recently using hyper spectral observation (Ben- Dor et al., 2001; Herold et al., 2003; Herold et al., 2004). The characteristic scale of urban reflectance is generally between 10 and 20 meters comparable to the Ground Instantaneous Field of View of most operational remote sensing sensors (Small, 2003). This author reported that if the GIFOV is smaller than the scale of the features, then they will be adequately sampled and thereby recognizable in the image; but if the GIFOV is comparable to the scale of the features, then the measured radiance will generally result from a combination of surfaces of different reflectance within the GIFOV. Moreover,optical sensors are said to have limited swath widths and spatial resolution. As a consequence, the measured radiance is a result of both the Ground Instantaneous Field of View and the scale of the feature in the urban mosaic (Small, 2001). Hepner and Chen (2002) used the optimal spectral analysis to distinguish the different urban materials in the urban area of Park City, Utah using AVIRIS data. The research applied the Spectral Angle mapper method to discriminate the different urban materials at two different scales. The Spectral Angle Mapper method showed that healthy grass and dry grass can spectrally be discriminated as well as turbid and clear water. In addition, four different roofing materials namely asphalt based roof, metal roof, wood shingle, membrane were discriminated.

Several other methods for urban spectra investigation have been proposed in literatures and among them are the Principal Component Analysis,the Bhattacharya distance,the Mahalanobis distance (Narendra & Fukunaga,1977) ,the Discriminant analysis feature extraction method (Duda & Hart,1973)the canonical correspondence analysis(Ter Braak & Prentice,1988) , the discriminant analysis (Tabachnick & Fidell, 2001) and the minimum distance feature space optimization.

In order to evaluate the spectral separability of different urban materials, Weng and Quattrochi (2007) compared the reflectance of several urban materials according to the wavelength. The study used the Bhattacharya distance approach which is a measure of spectral separability to evaluate the extent to which various urban materials can be spectrally distinguished. As results they found a prominent absorption of visible light and NIR radiation by red tile roofs attributed to the presence of iron oxide. In addition, the reflectance of fired bricks has been found to be higher towards longer wavelengths, because of the loss of water in the production firing process. Furthermore it was revealed that there is a difference in terms of reflectance between new asphalt and the old one. New asphalt reflectance was found lower in short wavelengths and higher in long wavelengths. In the same way the reflectance of parking areas was revealed to be higher in short wavelengths and lower in long wavelengths.Moreover, Green residential grass reflectance has been found higher between 650 and 800 nm and lower for wavelengths longer than 800 nm. Also, green vegetation on old asphalt have shown a reflectance varying between 0 and 8 for wavelengths shorter than 650 nm, for those longer than 650 nm the reflectance varied between 10 and 15 (Weng & Quattrochi, 2007).

Furthermore,the research identified a set of optimal wavelengths appropriate for urban mapping. The study suggested the following bands: 400 to 800 nm, 1600 to 1800 nm and 2100 to about 2300 nm. However,a certain number of urban materials was found difficult to separate. For instance, asphalt roads were found difficult to separate from parking plots, while grey tile roofs were found difficult to separate from composite shingle and tar roofs. Equally important, asphalt roads were difficult to separate from composite shingle, tar, and grey-tile roofs. Also bare soil surfaces were found difficult to distinguish from concrete roads (Weng & Quattrochi, 2007).

Another technique of spectral separability tested in researches is the Principal Component Analysis. The method reduces the feature dimensionality, by maximizing data variance in a reduced number of selected features. But, the approach does not necessarily improve object information quality within the reduced number of selected features. Moreover, the Discriminant Analysis feature extraction method was reported by Duda and Hart (1973) to be limited in discriminating very close classes such as roads and roofs and also produces a limited number of high level features up to a number of class minus one. Additional techniques of object spectral separability can be found in Jimenez & Landgrebe (1999).

2.3.4 Measures of object context

In object-based classification, more and more researchers are looking beyond low-level colour, size, texture and shape features for more effective objects extraction. In fact, a major limitation of individual material detectors is the significant number of misclassification that occur because of the similarities in colour, texture, size and shape characteristics (Singhal et al., 2003; Bruzzone & Carlin, 2006). It is widely acknowledged that correctly recognizing isolated patches of pure materials without context is extremely challenging, even for human observers (Singhal et al., 2003). Context can be defined as any significant information that can influence the way a scene and the object within it can be perceived (Marques et al., 2001). Context information can come from an overall description of the whole scene or from the normal relationships among the locations of different objects or regions within the scene. In the case of a scene context, knowledge of an urban scene for example can reveal what urban material types may be present in the scene, and where these materials may be likely to occur.

Although the advantages offered by the use of spatial context are weaker than those made available by the use of scene context, the merits of spatial context are still sufficient to reduce the ambiguity among conflicting objects characteristics and eliminating improbable spatial configuration in objects detection (Singhal et al., 2003). Moreover, the existence of two types of spatial contextual relationships in images is reported in some literatures (Singhal et al., 2003). First, relationships exist between co-occurrence of certain materials. Second, relationships exist between spatial locations of materials or objects in an image. However, several spatial relations can be defined between objects within a scene and among them are distance relations, topological relations and direction relations.

Distance relation between objects is a very important knowledge for object recognition. In an aerial image context for example, a rectangular unknown object of small size near a house and located in the proximity of a road network is likely to be car in a parking area. On other hand, the same object located within a forest extent and away from a road network is likely to be an isolated small building. Similarly, a rectangular object in a house would be very difficult to describe without context. By associating a distance relationship with chairs and the room walls ,one could be able to determine if the object is likely to be a table or not. Context based object recognition uses known objects and associates spatial relationships to describe unknown objects.

In addition to distance relationships,object context can be define through topological relationships. Topological relations do not change under topological transformation and translation (Liu et al., 2008). As example of topological relations are meet, disjoint, contains, covers, equal relations.

Directional relationships define the direction of unknown objects in relations to an known object. These relations are preserved under translation and scaling (Liu et al., 2008). Examples of directional relations are left,right, above, below, underneath.

Fuzzy relations are vague estimations of distance relations between different objects . A few examples of fuzzy relations are near to, far from, close to.

Measures of objects spatial arrangement

The computation of building orientation is among many techniques that can be used to determine whether a group of building constitute a residential or an informal settlement area. But the methods investigated to compute object orientation are based on complex mathematical functions that are difficult to implement. In contrast, techniques used to determine spatial arrangements of patterns make use of easy statistical and mathematical components that can easily be implemented. However, several techniques and algorithms that enable the description of data spatial structure have been tested by many scholars and in a diversity of fields. The most commonly used methods for spatial structure investigation are the following:

1. The Quadrant Count Method.
2. The Joint-Count Statistic method.
3. The Nearest Neighbour Distance method.

Quadrant Count-Method(QCM) is a point spatial analysis techniques, that segments the data set into region of equal size called quadrants. Within each quadrant is counted the number of objects of interest . The method considers the distribution of points of interest within each quadrant as an indicator of the object spatial arrangement. To illustrate the method a square space was created in which 25 points were plotted .The data set was then divided into four quadrants to facilitate the understanding of the technique as illustrated by the figure 2.1.

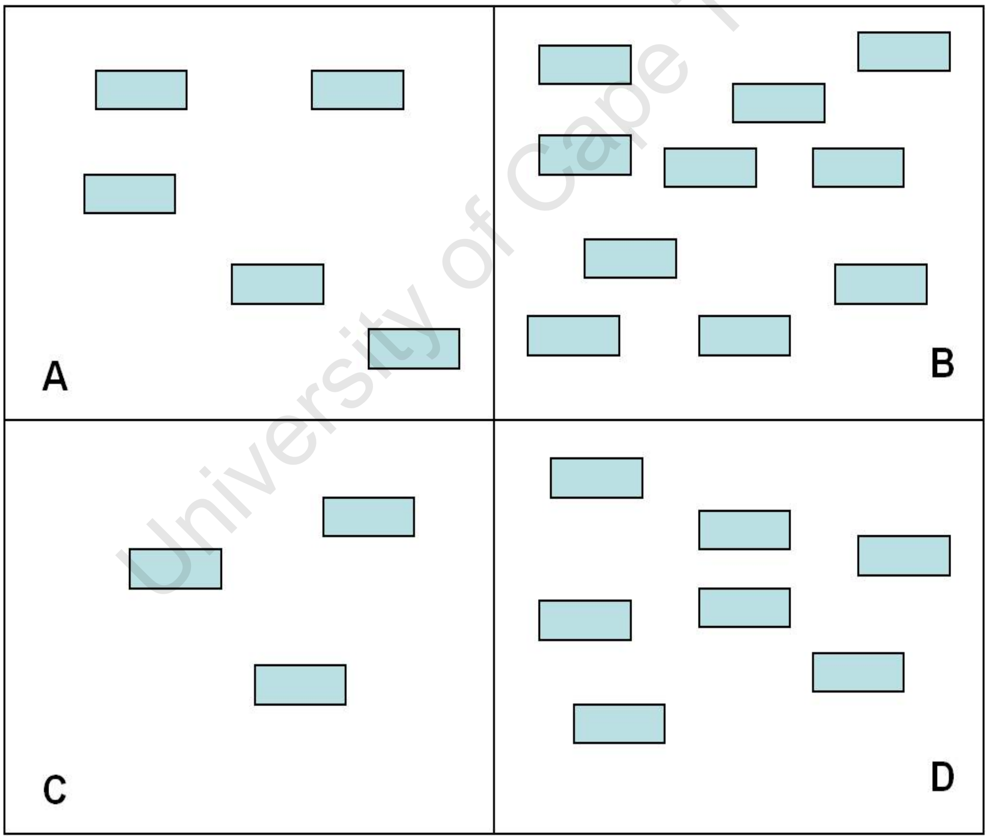

Figure 2.1: An illustration of a data divided into four quadrants: A and C show large areas without objects and B and D show cluster organizations of objects.

Once the segmentation is done, the Mean of objects of interest is calculated through the following equation :

$$Mean = \frac{objects\, count}{Quadrant\, count} = \frac{25}{4} = 3.75 \tag{2.10}$$

By considering x_i as the frequency of object of interest within each quadrant and n as the total number of quadrants,the variance of the objects is calculated as follows :

$$Variance = \frac{\sum x_i^2 - \frac{\sum x_i^2}{n})}{n-1} \tag{2.11}$$

The Variance To Mean Ratio can then be calculated as follows :

$$Variance\, To\, mean\, ratio = \frac{Variance}{Mean} = \frac{8.92}{3.75} = 2.38 \tag{2.12}$$

Since the Variance To mean ratio is greater than 1 from our illustration, our data set has one or more group of objects in clusters and large region within the data set are without objects .If the Variance To Mean ratio value was 1, then the object of interest would have been random, meaning that object distribution could have not dominant trends towards clustering or dispersion. On the other hand, if the Variance To Mean ratio was smaller than 1, the objects would have been regularly dispersed.

The basic of the join count statistics method is that the values of observations, which might be points or building at their respective locations, are given a value code which might be 1 for black colour and 0 for white colour.A map using this technique can be represented as a succession of black and white areas.Neighbouring objects to the object sample are assessed based on their code.The technique classifies joins between adjacent regions as either BB, WW,WB where BB represent a join between two black regions and WW a join between two adjacent white regions and WB a join between a white and black regions.

Moreover, many formulas have been made available through literatures to compute the Nearest Neighbour distribution index as follows :

$$Nearest\ Neighbour\ Distribution = \frac{Nearest\ Neighbour\ Mean\ Distance}{0.5 \times \sqrt{\frac{A}{N}}} \tag{2.13}$$

With A as the area and N the number of building within the study area.

Another equation that can be used to compute the Nearest neighbour distribution is defined as :

$$Nearest\ Neighbour\ distribution = 2 \times \frac{Mean\ Nearest\ Neighbour\ distance}{\sqrt{\frac{N}{A}}} \tag{2.14}$$

Objects distribution can also be calculated through a combination of three steps where the Mean Nearest Neighbour distance is computed a first stage and the Nearest neighbour Mean distance at Random using the equation:

$$Nearest\ Neighbour\ distance\ at\ Random = \frac{1}{2 \times \sqrt{Density}} \tag{2.15}$$

Where

$$Density = \frac{Number\ of\ building}{Area} \tag{2.16}$$

Once the Mean Nearest Neighbour distance and the Mean Nearest neighbour distance at random are calculated ,the standardized nearest neighbour index NNI is calculated to describe building distribution as follows :

$$NNI = \frac{Mean\ Nearest\ Neighbour\ Distance}{Mean\ Nearest\ Neighbour\ Distance_R} \tag{2.17}$$

2.4 Review of some urban classification studies

Several classification studies have been applied to urban areas which are composed of a diversity of land cover classes, such as roads, buildings, vegetation, parking plots and open space. Depending on the object of interest,urban classification can focus on the extraction of general land covers, such as impervious surfaces or built up surfaces, or can focus on more detailed subclasses such as roads, buildings, trees, water, open space. This section investigates some classification studies used to extract detailed urban subclasses. Among them are road extraction techniques (Hinz & Baumgartner,2000; Peteri & Ranchin,2003; Marangoz et al.,2004; Lee,2005; Zhang et al.,2009), open space extraction techniques (Maktav et al.,2001), urban water extraction techniques (Yue et al.,2008), urban vegetation extraction techniques (Shen et al.,2004; Alejandra et al.,2004).

2.4.1 Investigation of road extraction studies

Three main techniques have been widely used to extract roads from aerial imagery : (1) the classification based ,(2) the model based and (3) the knowledge based techniques. The classification based technique classifies roads based on their physical properties such as size and shape.The model based technique models roads into several parts and classifies each part based on context and spectral properties. The knowledge based technique defines first generic knowledge of roads and classifies them based on this knowledge.

Lee (2005) tested a three steps knowledge based road extraction technique on an Ikonos image of JeJu area in Korea. In the first step of the analysis, the author defined masked areas by eliminating areas of low occurrence of roads such as dense vegetation area. The second step of the technique defined road seed column from which the extraction originated, using a directional texture operator that took into consideration the advantages of linear properties of roads. The third step compared the spectral range of the seed column to surrounding pixels and merged to the seed column those that satisfied the minimum heterogeneity condition. The approach produced interesting results and had the advantage of limiting human intervention in the process, as the extraction was done automatically. Even though, the approach successfully extracted the targeted roads, the author reported some misclassified areas and connecting gaps between some road segments.

Moreover,Zhang et al., (2009) tested a technique that classified roads, based on their spectral, size, shape properties and context information. The authors combined all the road information in a semantic model on which the extraction was based. The quality of results was reported to have reached an accuracy of 85 percent for highways extraction, which was good. The extraction of main roads produced less accurate results illustrated by a classification accuracy of 77 percent. However, the authors acknowledged the interference of objects such as trees and buildings, that altered the quality of road extracted roads.Further knowledge on this technique can be found in Marangoz et al., (2004).

Roads had also been classified based on predefined models. Hinz and Baumgartner(2000) tested a technique that used a model comprising road radiometric, geometric and topology information. The approach used was based on two stages, in the first stage, the authors generated road global context information using down sampled image information. A digital surface model was used in this stage to define road height context relations. The second stage focussed on the construction of road lane segments, based on the grouping of road markings. The research reported good results improved by the use of height information that enabled the separation of roads from buildings.Similar studies were done by Petiri and Ranchin, who used road models to extract road from Quickbird and Ikonos images.

2.4.2 Urban water extraction techniques

Water extraction in urban areas has been tested by many scholars in the last decade (Zhang et al.,2004). More recently, Yue et al.,(2008) extracted a section of Chaping river in Mianyang city in China from SPOT5 imagery. The authors used a three steps object oriented approach to classify the river. The first step mainly consisted of data pre-processing in which, the imagery was projected, re-sampled and co-registered, in order to produce a more meaningful hybrid SPOT image, resulting from the fusion of a panchromatic and a multi-spectral image. In order to preserve the high spatial properties, the 10 m resolution multi-spectral image was co-registered to the 2.5 m panchromatic image. Principal component analysis technique was used to merge the two images. The authors used edge based segmentation technique because of its advantage of requiring only the scale parameter as segmentation parameter. In order to prepare a successful classification,the authors computed water spectral and spatial attributes, such as size, shape and the classification overall accuracy found was excellent and reached the value of 95.5 percent. Some limitations of the study were reported, as some part of Chaping river were merged with some parts of the road located along the river.

2.4.3 Urban open space extraction techniques

Few studies have focused on the extraction of urban open space from aerial imagery. The more recent study done by Maktav et al., (2011) tested a technique in order to extract potential open space from Quickbird images of Berlin and Ruhr in Germany, and Istanbul in Turkey. The technique used an object-oriented approach that classified at a first stage the basic land cover classes such as built up, trees, shrubs, grass, mixed vegetation, bare soil and water. At a second stage the authors merged the first stage classification results to produce four main classes namely settlements, forest, semi natural and water. After generation of the four main classes, their density maps were computed in order to identify non built up areas. The next step after the identification of non built up areas was filtering. The filtering of the non built up areas was based on shape and size thresholds in order to remove all small regions not suitable for construction purposes. The overall accuracies found for the three cities were above 80 percent. Open spaces in Ruhr area were classified with an accuracy of 87 percent and open spaces in Istanbul reached an accuracy of 90 percent. In Berlin, open spaces were found with a classification accuracy of 84 percent. Although the results were satisfactory in general, an examination of the different error matrices provided by the authors revealed some important confusions between open space and other classes. For example in Ruhr region, 1660 open space pixels were classified as built up areas and 44 classified as trees and shrubs. In Istanbul area, 2501 open space pixels were classified as built up areas and 3339 were classified as grass land and mixed vegetation. In Berlin area, the confusion was minimized and this could originate from the fact that the authors used a very small number of sample points, compared to the two previous areas.

2.4.4 Urban vegetation extraction techniques

Several scholars have investigated the classification of green space within urban areas (Alejandra et al., 2004). More recently Shen et al., (2010) attempted the extraction of urban vegetation from Quickbird imagery of Xinzhuang area in Shanghai,China. The authors tested three classification approaches namely Decision Tree, Maximum Likelihood and Minimum Distance. Before applying any classification algorithm on the data, the authors investigated the spectral properties of green space in the four bands of Quickbird image.The investigation revealed that green vegetation reflectance in band 4 is greater than the reflectance in band 2 which is greater than the reflectance in band 1, which is greater than the reflectance in band 3. The extraction of vegetation was also supported by the use of Normalized Difference Vegetation Index(NDVI) to separate vegetation from other urban classes. The results found revealed the superiority of the Decision Tree technique over the Maximum Likelihood and Minimum Distance techniques, with respective overall accuracies of 95.31 percent, 85.94 percent and 75 percent.

2.5 Investigation of some post classification tools

Pixel based classification techniques produce results with lower accuracy compared to object based classification techniques. However, the basic unit in the construction of image objects or segments remain pixel values.So there is still a need for post processing tools in order to improve classification results. Several tools have been made available within some commercial software such as Erdas imagine, Envi or e cognition. Erdas imagine provides a wide range of post classification filters such as the median filter, useful to correct the effect of salt and pepper caused by mixed pixels. In addition, ancillary data such as building density, can be used to improve classification results. In fact, high residential and commercial buildings have similar spectral response but considerably differ in density.

Other techniques such as polygon reshaping, manual merge and manual classification tools are available in the commercial software e cognition, can be useful to improve classification results.

2.6 Conclusion

This literature review was done in order to evaluate the most recent works done in multi-scale classification,as well as techniques of scene understanding, which are the main focuses of this research. This involve the investigation of the most used segmentation and classification techniques, techniques describing different urban objects and the different spatial relations defining them within the scene. Nine of the most advanced segmentation techniques were investigated in this chapter, but the choice of segmentation technique relies on the type of data to be used, and the type of object to be extracted(Dey et al.,2010). Based on this investigation and the previous work done by Dey et al., (2010), coarse scale segmentation are suitable for land cover mapping projects and in such cases, clustering segmentations techniques would be suggested. On the other hand if the purpose of the project is to classify land use classes, finer scale segmentation techniques would be the best and here multi resolution techniques could be more useful. If the four best techniques among the nine investigated in this chapter have to be ranked based on the quality of results, the time cost spent for processing, the simplicity and robustness of the method, the multi-resolution segmentation techniques would be selected as the best choice. The techniques integrates all other parameters taken into account by most of other segmentation techniques such as texture, pixel reflectance, size, shape, and topology. For images with a high texture component, segmentation based on Markov Random Field approach is suitable. Fuzzy approaches is suitable for very complex regions were the detection of object edges may create confusions. And finally the techniques based on the Neural Net work would is appropriate if the project does not require high quality segments.

In this chapter five main groups of the mostly used and advanced classification techniques ranging from pixel-based to per field based approaches were investigated, as well as techniques integrating context and prior knowledge. Per pixel classification has displayed limitations in complex areas such as urban areas, when used with high spatial resolution data. To deal with these limitations, sub-pixel classification techniques have been employed and proven to produce better results when used with coarse and medium resolution data such as Landsat images. Per field classification techniques on the other hand, have proven high quality results when used with high resolution images such as SPOT data, Ikonos, Quickbird or aerial photographs. The technique reduces the influence of spectral variability within certain land covers, by considering groups of homogeneous pixels instead of single pixel. Moreover, the use of additional context information has been reported to improve the classification results, as for instance a swimming pool and a agricultural Dam which are both water bodies, with similar spectral response, but can be separated based on their local context within the scene. In addition, knowledge-based techniques have been reported suitable to handle external data or knowledge to improve the classification quality.

In order to choose the best classification technique, a detail investigation of the different methods is required and the goal of the related section was to evaluate the best classification techniques. Based on classification results, processing time, quality of information integrated into the classification, the following three methods could be classified as the best : per field classification approaches with special stress on the hierarchical object-oriented classification technique, the context-based classification techniques and finally the knowledge-based approaches.

Moreover,the scene description investigation revealed that the area measurement remains the most used technique to estimate objects size. Area measure can be calculated from polygon outlines and the accuracy of the measure strongly relies on the accuracy of the object outlines.The investigation of different techniques of shape estimation revealed that the best techniques would be the perimeter to area ratio techniques and the parameters of related circles as they are easy to compute.The direct comparison techniques produced very interesting results but have the limitation of depending on the accuracy of the object outlines drawn by the analyst and is suitable for large objects such as villages and town outlines. The dispersion of elements of area techniques take into account area of mismatch between the shape under investigation and the standard model but are difficult to compute when dealing with very large and very small objects such as buildings.

At the exception of the principal component analysis method, all the spectral separability techniques investigated have been reported to produce interesting objects spectral separability. However, the choice of the appropriate technique can be influenced by the experience of the analyst with the technique. The minimum distance feature space optimization tool available in the commercial software e-cognition,the Bhattacharya distance techniques would be the best techniques as they take into account all the high level spectral features selectable by the analysis, compared to methods such as discriminant analysis feature extraction or the canonical correspondence analysis.

Moreover, among the context estimation techniques investigated in this chapter, the distance relation would be the more accurate compared to fuzzy or orientation relations. One advantage of the distance relation is that it can easily be computed in most of GIS tools available and can be used for point or polygon data.

From the investigation of some recent urban classification studies done in this chapter, knowledge-based classification technique was revealed the best for road extraction as the approach combines spectral, size, shape and context information. Among the two knowledge-based techniques proposed,the one used by Zhang et al.,(2009) was the simplest and produced very interesting results. The classification based techniques use limited information, which makes the extraction more challenging in complex areas such as urban areas. The model-based approaches are not always easy to implement as the information such as road lane are not always easy to extract from aerial imagery. The object-oriented approach used to extract water was revealed suitable as the integration of information such as size, shape and context made available by the technique contribute to the amelioration of classification results as reported in Yue et al.,(2008). The technique proposed by Markav et al.,(2011) remains one of the best if the purpose of the study is to identify potential areas for urban construction. Other techniques such as extracting the area measure of built up, vegetation, water from the total area of the study site can give an estimation of open space available within the area. Moreover, techniques combining spectral information and NDVI have proven to be efficient for the extraction of urban vegetation. The use of additional information such as object size, shape and context can improve vegetation extraction in urban areas.

It is very difficult to determine which technique is the best post classification approach to use as it depend on the nature of the project and the algorithm used. For traditional classifiers, median filters can be suggested to reduce the salt and pepper effects. For object based classification, rule based reclassification or manual reclassification can be suggested. These results were relevant for the object of this research as it gave an update of the different researches done in image classification as well as the latest developments in terms of software. From these, the multi-scale segmentation approach and the object-based classification technique would be suitable for this study because the combination of these two techniques gives potential to derive and exploit objects contexts, which is very relevant for this research.

Chapter 3

urban scene contextual model for multi-scale classification

3.1 Introduction

While the development of new high resolution imagery brings excitements to the field of urban analysis, some challenges still remain. Remote sensing records what is on the land surface, whereas in urban analysis it is land use that is of main interest (Mesev, 2007). Remote sensing only provides information at land cover level and linking land cover to land use is a great challenge. Land use refers to the function attributed to a land, for instance, a vegetated land can be of agricultural use and is characterized by certain attributes such as its proximity to water bodies, its location in respect to residential areas as well as certain shape, size, spectral ,textural characteristics. From this illustration, it appears that object recognition within a scene requires certain descriptive information about the object or a model without which, its accurate identification is difficult because object recognition does not happen in isolation and is influenced by the presence of other surrounding objects, as well as the overall scene context. Context information can be used in image classification in order to improve object discrimination within a scene(Schumm,1963; Boyce & Clark,1964; Lee & Salle,1970; Cressie,1993; Maceachren,1985; Chou,1995; Kumar et al.,2008; Williams & Wentz,2008; Liu et al.,2008; Bose & Grimson,2004).

In fact, high level features such as shape, size; spatial, spectral context can capture object variability and improve their discrimination within a scene. For instance, the distinction between a pedestrian and a car on a road is difficult without context. The separation between the two objects is possible not only based on pixels inside the object outlines, but also relies on many other parameters such as the surface on which the person or car is standing, the orientation of the viewer, the texture, the size, shape, the colour of the objects, as well as their spatial relations in reference to other neighbouring objects. All these parameters involved in the recognition process can be grouped into five categories: (1) the semantic context which defines the likelihood of some objects being found in some scenes but not in others, (2) the spatial context or positional context which defines the likelihood of finding some objects in certain positions and not others in respect to other objects in the scene, (3) the scale context or size context which defines the likelihood of some objects to have certain size and not others in respect to other objects within the scene , (4) the shape context which defines the likelihood of some objects to hold certain shape signature and not others in respect to other objects within the scene and (5) the spectral context which defines the likelihood of certain object types to produce certain spectral characteristics in certain spectral bands and not others in relation to other objects.

The goal of this chapter is to build an urban scene contextual model that will enable to train classification algorithms and improve classification accuracy on urban areas. That formulation of the model starts with a selection of land use samples based on a visual analysis of the scenes. The selection of objects samples from the different chosen land uses will enable the selection of high level features characterizing the objects. These high level features include object size features such as area measure or perimeter, shape statistical features such as shape compactness, mean shape measure, shape compactness standard deviation, pixel signature features such as mean value in the red, green, blue bands, mean standard deviation value in the red, green and blue bands and spatial relations such as distance building-trees, building-roads, building-building.

3.2 Material and Methods

3.2.1 Study area and Data

Located at Latitude 33° 55' 0" South and Longitude 18° 25' 0" East, Cape Town is the second-most populous city in South Africa and the provincial capital of the Western Cape. The city population is estimated at approximately 3.5 millions of people but the recent population census projected the population to be 4 255 857 by 2031. Besides being the legislative capital of South Africa and the provincial capital of the Western Cape Province, Cape

Town is one of Africa's most extraordinary tourists destinations. Its harbour and biodiversity contributed to make the city one of the most beautiful in the continent. The region covers an area surface of 2,500 square kilometres comprising six suburb groups namely the City Bowl, the Northern suburbs, Eastern Suburbs, the Atlantic Seaboard, the Southern Suburbs, the South Peninsula and the West Coast suburbs(Figure.3.1). The area covering City Bowl suburbs have been selected for this investigation due to its variety in land use and land covers including open space, grass land, tree plantations, recreational parks, sport fields, roads, residential and non residential buildings, parking plots and water bodies (Figure 3.1). The city Bowl area is bordered by Table Bay and defined by the mountains of Signal Hill, Lion's Head, and Table Mountain. It comprises the central business district of Cape Town, the harbour, the Company's Garden, and the residential suburbs of De Waterkant, Devil's Peak, District Six, Zonnebloem, Gardens, Bo-Kaap, Higgovale, Oranjezicht, Schotsche Kloof, Tamboerskloof, University Estate, Vredelhoek, Walmer Estate and Woodstock.

High resolution satellite data would be the best for urban analysis, but due to the high cost of the data, the choice was oriented towards an aerial photograph. This remote sensing image is a mosaic of 10 aerial photographs acquired in 2009 and covering the area of Cape Town. The image mosaic was already created before collection from the Geomatics Department, University of Cape Town. Global Mapper and Ortho Vista software enable the creation of image mosaics by importing all the images into the software and then exporting them as a merged raster file. The 0.5 m spatial resolution image possesses three spectral bands covering the visible light spectrum composed of Red, Green, Blue channels. Due to the large extent of the area, a subset was created compression free in order to preserve the spatial and spectral properties of the data, as data compression alters the quality of the data (Campbell, 2006).The figure 3.1 bellow shows the extent of the study area.

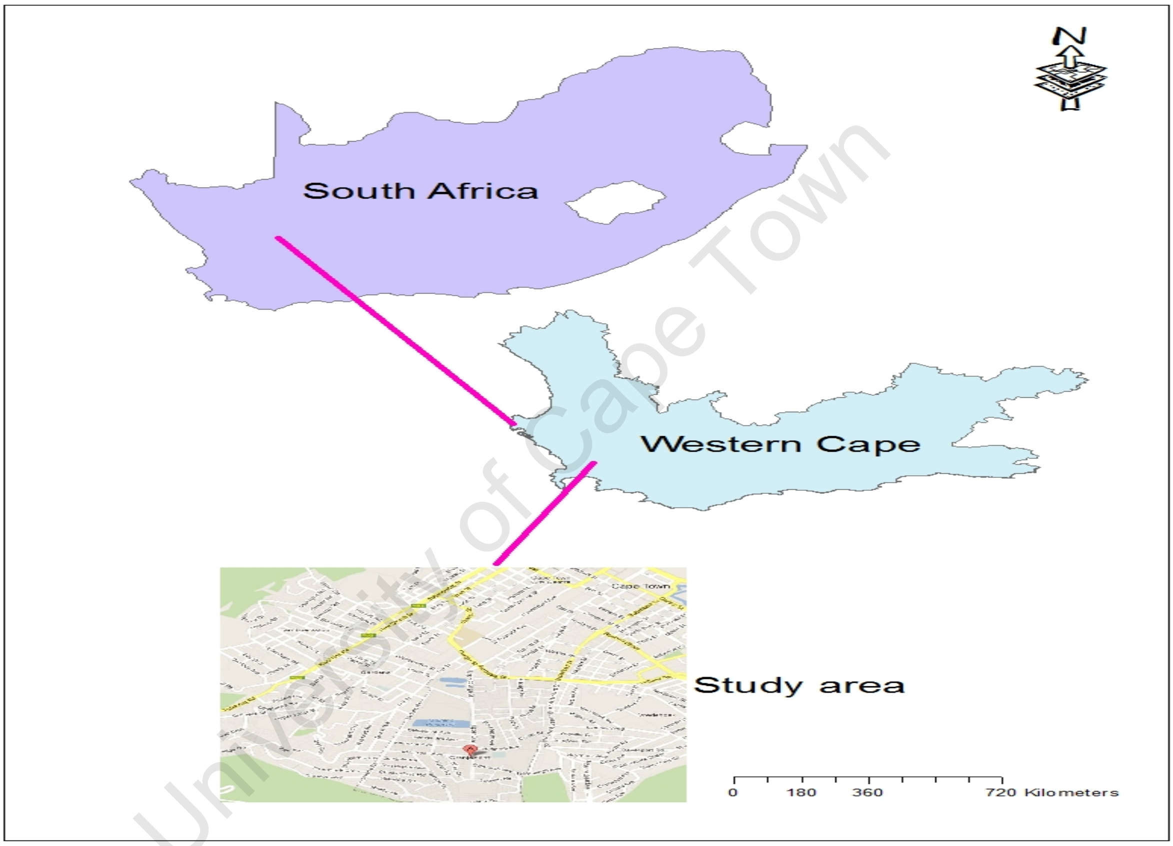

Figure 3.1: The study site of Cape Town including the different suburbs falling into the area

The vector data used in this research comprise polygon outlines digitized from the remote sensing image. The different outlines were digitized in order to compute objects properties such as area and perimeter measures, shape compactness, spatial distances, as well as objects spatial organization index. Polygon outline digitizing remains the most accurate but tedious technique for object geometric characteristics extraction. When digitizing the different objects outlines, snapping distance of 2 was chosen as larger distance can alter object shapes (Chang, 2008).

Figure 3.2 summarizes the different methods used in this investigation. Visual analysis of the image involves the identification of different land cover classes within the scene. Once identified, the land cover classes are digitized in order to extract outlines of their distinguishing objects. High level features such as area, perimeter measures, shape compactness, spatial arrangement and separating distances are derived. The feature space optimization enables the identification of suitable high level spectral features that minimize spectral overlap between classes by calculation the spectral distances between classes in a spectral feature space. The most relevant features that minimize overlap among classes are selected to build a scene model that will be used for urban scene understanding and classification.

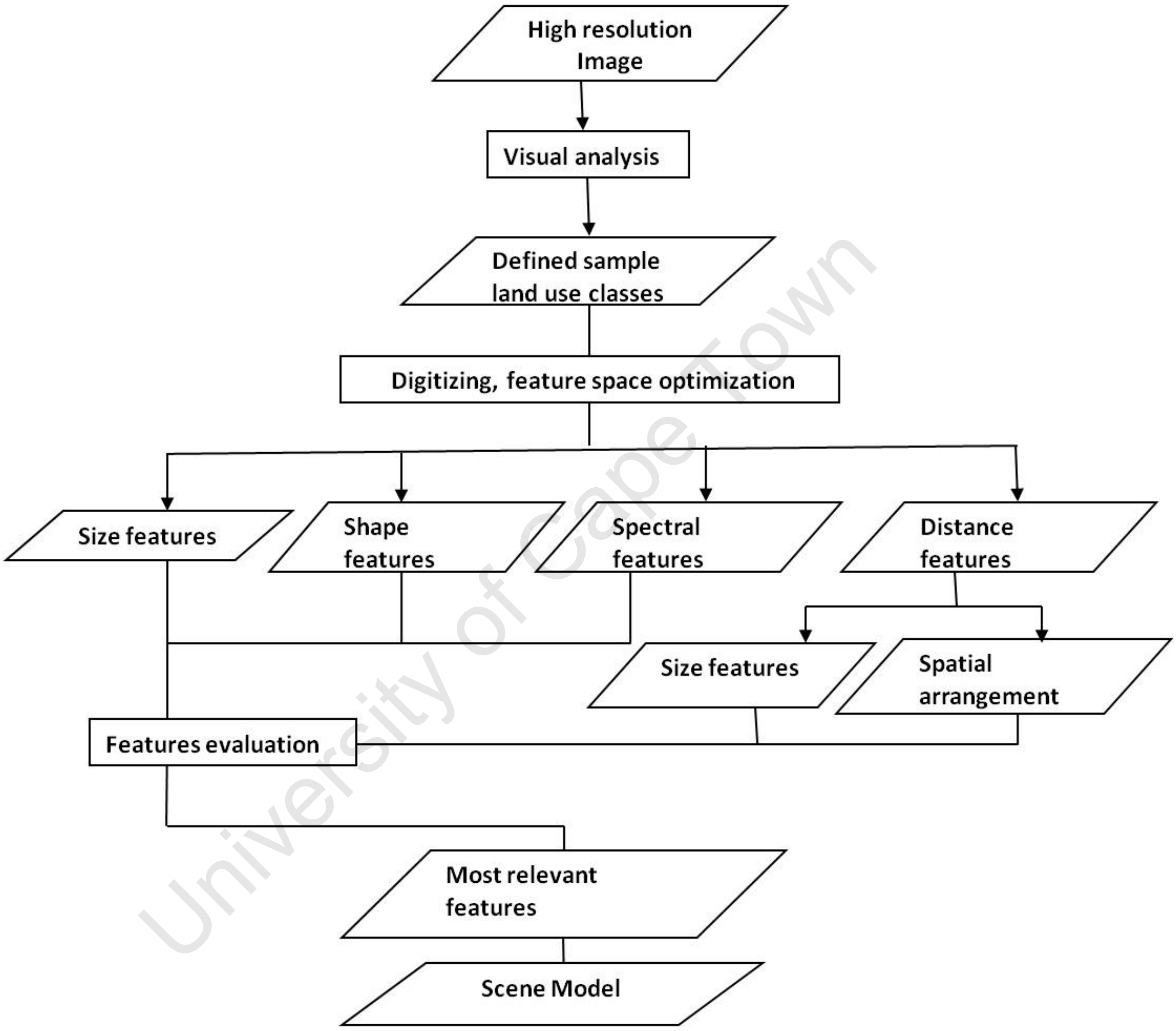

Figure 3.2: After a visual analysis of the high resolution image, a set of land use classes is selected in order to derive the different high level features which are evaluated in order to build a scene model.

3.2.2 Urban land use selection

A visual analysis of the study area enables the selection of seven main land use classes. These include all possible urban objects types so that when classifying an urban scene it is likely to fall into one of these categories. The number of land use classes was limited to thirteen as too many classes would lead to a redundancy of attributes between classes. The following were chosen:

1. Residential buildings.
2. commercial buildings.
3. Industrial buildings.
4. Educational buildings.
5. trees.
6. Roads.
7. Parking plots.
8. Water (dams).
9. Swimming pools.
10. Artificial sport fields.
11. Open space.
12. Recreation parks.
13. Grassland.

Figures 3.3 and 3.4 illustrate the land use samples selected for this study. The figure 3.3 presents four different land use and land cover classes that includes grassland class which is generally used as sport fields in urban areas. The size of sport fields vary from one area to the other, depending on the type of sport practised on. Most of sport areas exhibit regular shape closer to rectangular shapes, the spectral reflectance varies as well and sometimes overlaps with bare ground. Artificial surfaces presented in the figure are also used for sport purposes but are generally located around educational buildings. They have generally rectangular shape and reveal high reflectance in the red band and sometimes overlap with features such as parking or roads that share similar spectral properties. Non residential areas are mostly composed of elongated buildings with the shape varying from L to U models. The size of parking plots depend on the use attributed to them. Parking size can vary from one area to the other and are generally larger within commercial and industrial areas. Some residential areas can have large parking plots as illustrated in the figure 3.3. Their spectral characteristics vary from one band to the other but are generally high in the red band and share similarities with roads and some building roof types. Shrub class is generally located outside urban areas and exhibits spectral characteristics closer to bare ground.

The following subsections will study the various objects attributes as well as some techniques used to extract them from the image.

3.2.3 Size and shape features

Objects' size context was captured through object size measure and various methods are available to estimate objects spatial size. The size of each object was measured as equivalent to the area measure of its polygon outline. The different object outlines were manually digitized and the spatial area of each digitized polygon was calculated automatically from the calculate geometry tool in ArcGIS10. The measure unit was chosen as square meter as the targeted polygons were of small and medium size. Object size is a relevant information to support object-oriented classification approaches in urban areas, because of the high spectral variability, which negatively affects classification accuracies (Pozzi & Small, 2002). In fact, objects within the scene can be separated from each other based on size measure. Buildings and cars for instance, can be separated based on their respective size measurements. A car has an area size two to three times smaller than a building.

Moreover, there is a significant interest and need for shape analysis in urban mapping (Wentz, 2000). For instance, scientists have used shape indices to analyse spatial patterns in order to study urban growth. However, visually identifying and comparing regions on the basis of shape is easy and intuitive for humans to do, but difficult for artificial intelligence systems. As a result, numerous attempts have been made to quantify shape for the identification of spatial objects (Wentz, 2000). Objects can be identified based on shape features such as compactness and length or $\frac{Length}{Width}$ ratio (Williams & Wentz, 2008). The different object shape compactness values were calculated based on the equation 3.1.

$$Compactness = \frac{4 \times \mathit{\Pi} \times A}{P^2} \tag{3.1}$$

where A and P are respectively the polygon area measure and polygon perimeter. The ratio returns score values ranging from 0 to 1 where scores closer to 1 characterize more compact shapes and scores closer to 0 characterizing less compact shapes. Several methods are available to estimate shape measure but the perimeter to area defining shape compactness is a simple technique and can be implemented in most of GIS environments.

3.2.4 Extraction of objects spectral features

Object spectral context is very significant for accurately discriminating objects from one another within a spectral feature space. Spectral characteristics of objects can differ from one band to the other. For instance, ninety percent of spectral information characterizing vegetation is stored in near-infra red and red bands (Bonn & Escadafal,1996). Therefore, spectral bands can be considered as very relevant source of scene information(Khedam & Belhadj-Aissa, 2011). The technique used to spectrally separate objects from one another within a scene is called spectral separability analysis . The aim of such analysis is the identification of suitable high level spectral features that enable the separation of objects in a scene. Several techniques for spectral separability analysis have been investigated in the literature review chapter but for this study, the minimum distance feature space optimization technique was used. The minimum distance algorithm was trained by supervised method, relying on parametric signatures describing different land cover classes. Each training sample consisted of at least 120 pixels and satisfied the 10 n criterion, where n is the number of band to be used for classification(Kumar et al., 2008). The approach computed samples' mean spectral values and tested the different candidate pixels against those mean values based on the minimum spectral distance. Signatures were then assessed using histogram plots and error matrices, which present the different land cover classes, their respective samples and the different spectral distances separating them. After assessment of separability, spectral bands with good distances were selected to be used in further classification. Tables 3.1 presents respectively the different spectral bands and band combinations tested in this separability analysis

Table 3.1: Examples of high level spectral features considered in the investigation: on the left column are listed the spectral features and in the right column are listed their respective corresponding descriptions

Spectral Features	Feature descriptions
Mean Red	Mean pixel value in the red channel
Mean Green	Mean pixel value in the green channel
Mean Blue	Mean pixel value in the blue channel
Redness index	The ratio of the Mean red to the sum of Mean green and Mean blue
Greenness index	The ratio of the Mean green to the sum of Mean red and Mean blue
Blueness index	The ratio of the Mean blue to the sum of the Mean red and Mean green

3.2.5 Geometric and Topological Relations

If one say that a group of buildings is clustered, disperse, random or regularly arranged, the different buildings within the pattern are related by a spatial organization type but if a commercial building is located within a residential area, the two are related by a topology relation. There are many ways features can be related to each other within a scene. In this investigation, particular attention is given to geometric and topological relations. Geometrical relationships such as distance relations, change with the coordinate system. In fact, the distance between two points can vary from a flat map surface to an ellipsoidal earth surface. Topological relationships are not affected by the coordinate system. The fact that a commercial building is within a residential area or that buildings in a pattern are clustered, would not change whether you are looking at a flat map or an ellipsoidal earth.

An interesting approach for estimating distance relations is the Euclidean distance. Distance relations allow one to derive a complete collection of mutual proximity relationships for each combination of spatial objects. In the case of point data, respective map coordinates are used to calculate the distance between points. Objects' centroids were considered in this study in order to compute mutual distances between different polygons within the study area. Once the centroids identified, map coordinate were input in equation 3.2 to determine their respective Euclidean distances.

$$Distance = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{3.2}$$

With x_1, x_2, y_1, y_2 the map coordinates of two random spatial objects considered in the computation. The equation outputs distance estimations in meters as the different coordinates measures are in meters.

Understanding the spatial organization of objects composing a pattern, such as a group of buildings, is very important in order to describe the pattern. There are variety of techniques used to measure spatial arrangement of objects within a pattern and the mostly used and among the most accurate techniques, is the nearest neighbour dispersion index(Lee & Wong, 2001). The technique enables the identification of three types of spatial arrangements of objects, namely random, disperse and regular, based on a nearest neighbour dispersion scale. The figure 3.5 describes dispersion bar, used to classify spatial patterns using the nearest neighbour technique.

ArcGIS software(ESRI) was used to digitize different groups of objects and calculate the respective proximity distances separating them. The nearest neighbour index combines the mean nearest neighbour distance and the mean nearest neighbour distance at random, in a ratio, to produce the equation 3.3 that compute a score describing objects spatial organization within a pattern(see literature review chapter for more details).

$$NNI = \frac{Mean\ Nearest\ Neighbour\ Distance}{Mean\ Nearest\ Neighbour\ Distance_R} \tag{3.3}$$

The first variable representing the numerator of the ratio is the mean distance separating different objects within their respective pattern. This variable was calculated by averaging all the distances separating objects within their respective patterns. The second variable representing the denominator, was the mean distance separating objects if randomly distributed within the pattern they constitute. This was calculated by taking one half of the square root of the object density within the targeted pattern, as illustrated in the equation 3.4 and 3.5 below.

$$Nearest\ Neighbour\ distance\ at\ Random = \frac{1}{2 \times \sqrt{Density}} \tag{3.4}$$

Where

$$Density = \frac{Number\ of\ building}{Area} \tag{3.5}$$

The ratio produced scores varying between 0 and 2.4. The closest to zero is the score value, the more clustered are the objects within the pattern. An index of 1 characterized randomly distributed objects within the pattern. An illustration of such spatial organization is the spatial structure of informal settlement patterns. In contrast, regular arrangement of objects such as formal buildings within cadastral boundaries, are characterized by an index close or equal to 2.45 (Cressie,1993). The figure 3.6 illustrates two opposite examples of building organization.

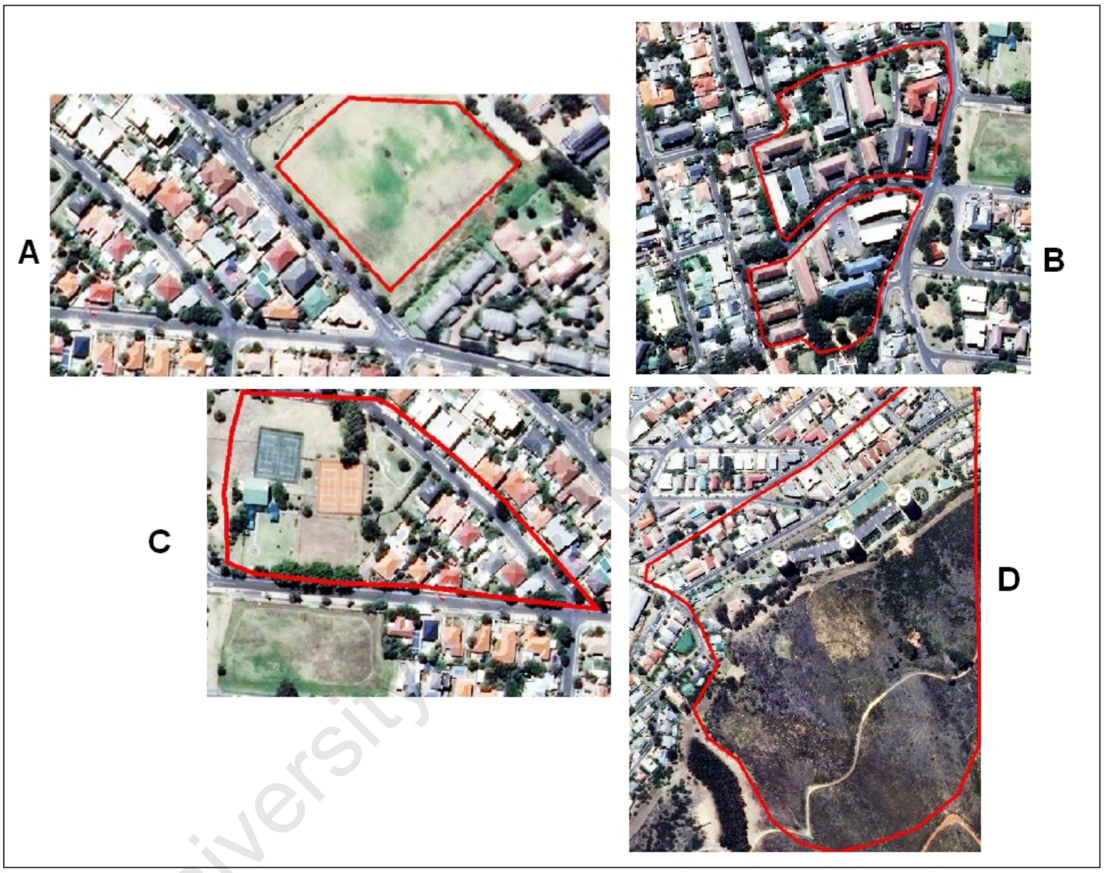

Figure 3.3: Figures A, B, C and D are three subsets of the study site illustrating the different land used samples selected in this investigation. Figure A shows an example of grassland, figure B shows examples of non residential buildings,figure C illustrates two examples of artificial sport fields and figure D reveals examples of parking areas, open space(bare ground), shrub and roads.

Figure 3.4: The polygon A on the left hand side image illustrates recreation parks, surrounded by trees and grassland. The two red polygons (B) on the top right image show examples of residential areas composed of small and medium size buildings of red, green and grey roofs. Polygon C on bottom right image illustrates a large dam(dark colour) surrounded by trees and buildings.

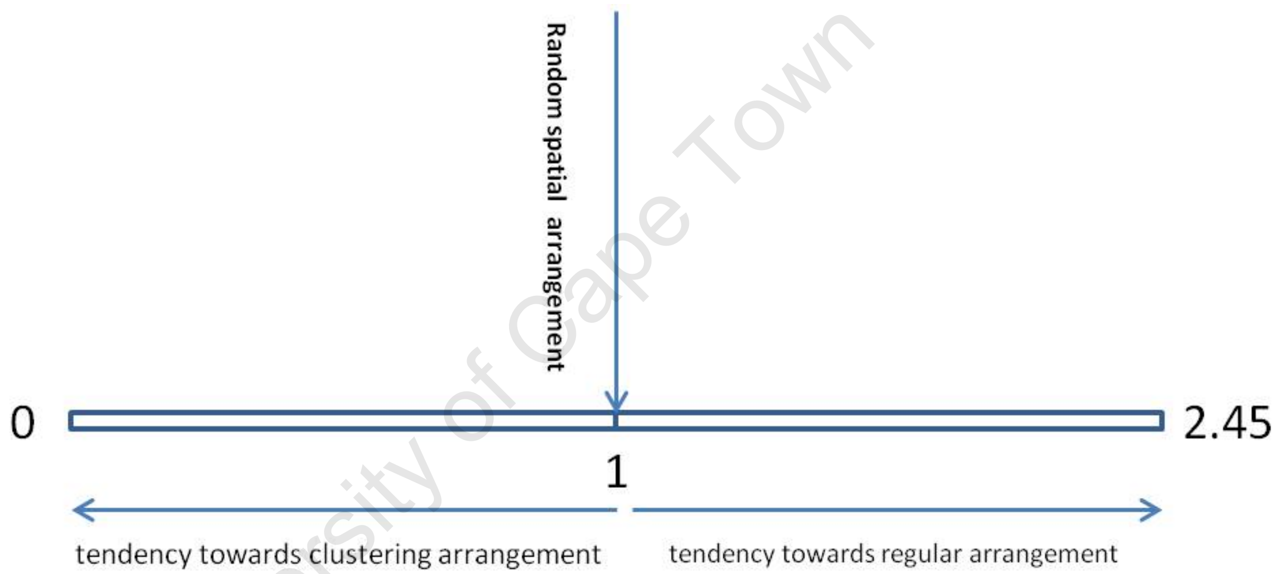

Figure 3.5: Spatial organization bar: Values closer to zero characterize clustered pattern,values closer to 1 describe random organization and values closer to 2.4 describe more regular patterns.

Figure 3.6: The scene A shows a more regular spatial organization of buildings, the space between buildings is more or less constant. The scene B shows clustered building with high variation of inter building distances.These types of patterns generally score a nearest neighbour index of 1 as buildings are randomly distributed.

3.3 Experimental results

In the first phase of selection of high level contextual features, the attention was directed to scale context analysis, since size signatures play a key role in improving urban settlements classification. Figure 3.7 compares the different size of urban settlements in the study area and, reveals the simplicity to separate them based on their respective size measures. Therefore, pure spectral signatures alone cannot effectively enable the separation of the different urban settlements, due to spectral similarities(Lu et al., 2010). Residential buildings dominate the entire urban settlement class and the size of buildings vary between $100m^2$ and $300m^2$. Commercial, education and industrial building are characterized by building size greater than $300m^2$. As the results demonstrate, the size of a building can indicate its use or function within the area. In addition to the variability of building size between settlement patterns, the size of buildings can also differ from one another within the same pattern.

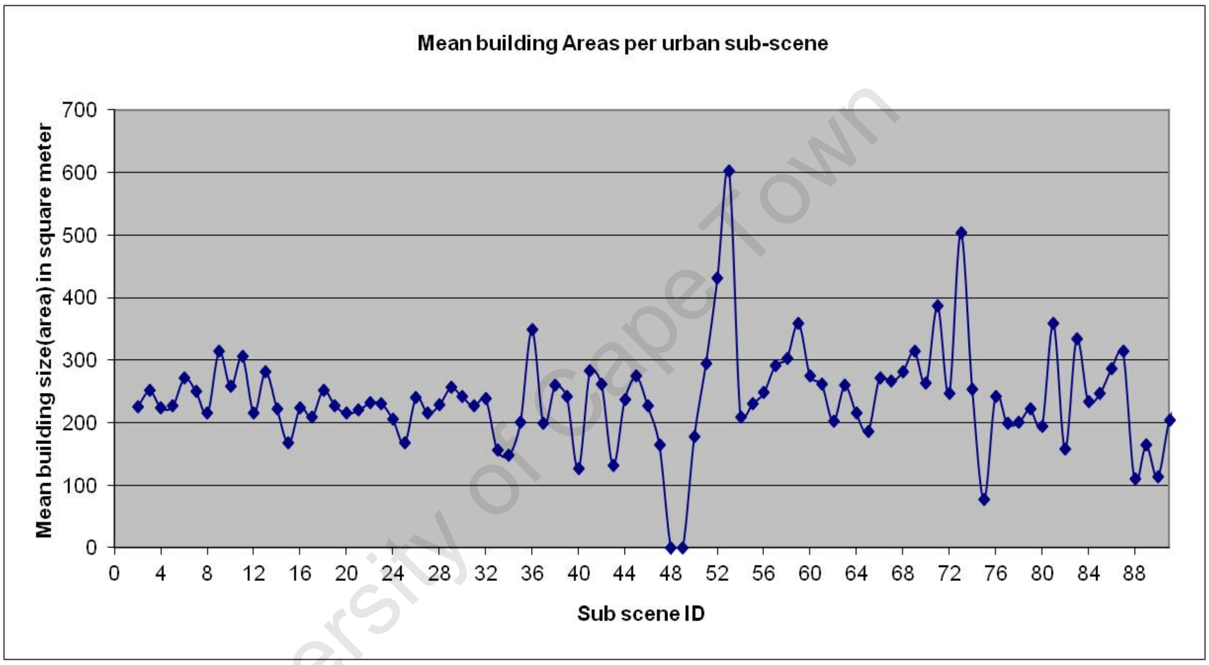

Figure 3.7: Variation of building sizes per area: sub scene 1 to 44 are examples of areas composed of residential buildings with the size varying between 100 and $300m^2$, whereas sub scenes 8,10,51,52,57,58,68,70,72,80,82 and 86 are examples of areas composed non residential buildings with the size greater than $300m^2$.

The figure 3.8 presents a visual illustration of a commercial and an educational building characterized by size measures greater than $300m^2$. The lengths and widths of buildings are respectively greater than 60 and 40 m,which are larger than normal residential settlements located in their respective neighbourhood.

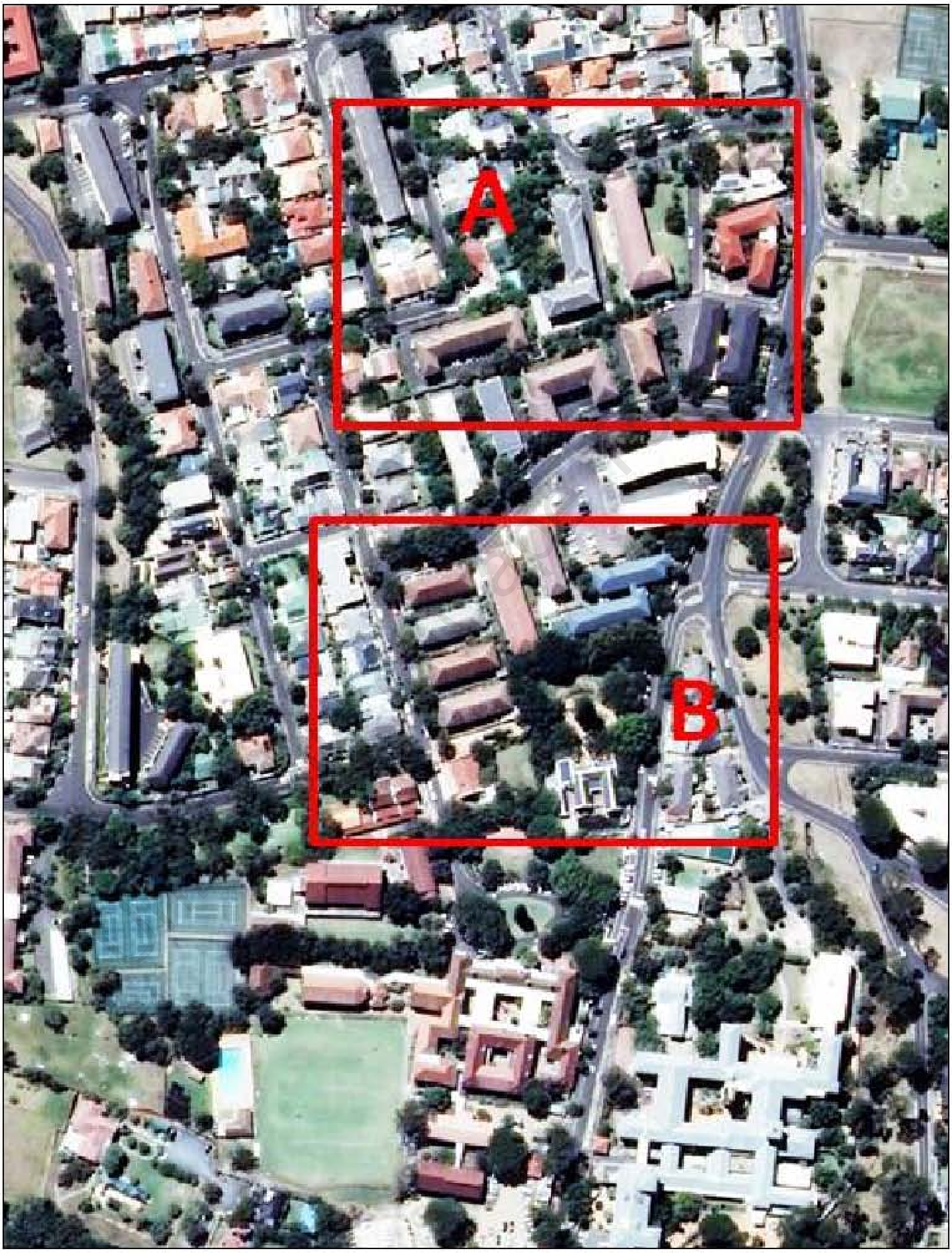

Figure 3.8: A and B :examples of Non residential buildings. From visual observation, the buildings have respective size larger than other surrounding buildings in the area.

Since roads incorporate significant spatial information,the extraction can rely on lengths, which distinguish them from other features as road are characterised by lengths greater than 70 m. Road discrimination can also rely on width measures as road widths vary from one urban area to the other. In low density areas, near urban peripheries, the width varies from 5 to 6 m, and in high density areas, towards the city centre, road width varies from 9 to 11 m. Around the mountains and natural vegetation areas, road width varies from 2 to 4 m.

Table 3.2 presents the results of shape signature analysis in the form of compactness values. The algorithm used in the study compares object polygon area to the smallest digitized circle that contains the object outlines. Compactness values close to 1 describe objects with shape closer to a circle and values closer to zero describe more elongated shapes. Tree plantations were found with shape outlines closer to circle with a compactness score greater than 0.70. In fact, vegetation generally shows shape heterogeneity as a result of irregular shadow or shade (Blaschke, 2010). Similar results were found for water bodies dominated by swimming pools mainly characterized by more irregular shapes. Urban settlements were found with more regular shapes described by a shape compactness value slightly above 0.5. Shape compactness values of 0.459 and 0.337 were found respectively for recreation areas (parks) and parking plots. Roads can be separated from other built up structures based on their very small shape compactness scores, approximately 0.049. A $\frac{Length}{Width} \geq 1.6$ ratio can also help to separate roads from other urban objects.

Figure 3.9 presents histogram plots showing spectral separability between the tree class and roads in the red and green bands. Most of the pure spectral information characterizing tree class are located in the spectral range from 15 to 83.8 in the red band. In this range the objects can easily be isolated from the remaining objects present in the scene.This band contains important spectral information characterizing building roofs in the spectral range from 100 to 255. The red and green bands contain relevant spectral information characterizing roads in the spectral range from 98 and 191.3, out of this range the feature may overlap with trees. In fact, the vegetation dominates the spectral scope from approximately 23 up to 94 in the green band.

Table 3.2: Compactness values per land cover type: Roads and parking plots are examples of irregular shaped objects whereas buildings are examples of more regular objects. Similar shape signatures are observed between trees and water bodies on one hand and between Recreation areas and Parking plots on the other hand.

land covers	mean land covers compactness values
buildings	0.623885
tree plantations	0.700957
parking Plots	0.336959
parks	0.459092
water bodies	0.708861
roads	0.0049361

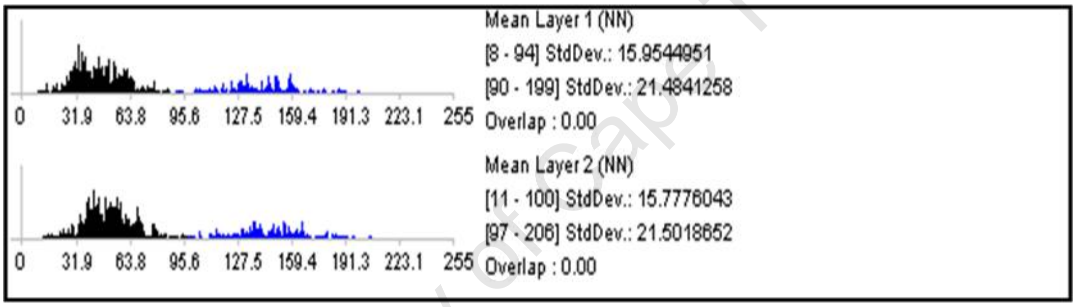

Figure 3.9: Histogram plots showing trees and road dominant spectral ranges in the red (Layer 1) and green(Layer 2) bands: Trees, in dark black colour dominate the spectral range between 15 and 83.8 in the red band and 23 to 94 in the green band.

Figure 3.10 confirms that important spectral information characterizing trees and green tiled roofs are contained in the red band and a spectral distance of 2.034 separates the two classes. However, a certain number of urban materials are difficult to spectrally separate. It was found that buildings are difficult to separate from parking plots, with respective spectral separability distance of 0.051. The small separability distance of 0.002 between grassland(sport fields) and parks makes the separation of both features very difficult. Furthermore, the separability distance of 0.904 between trees and parks complicates the isolation of both features.

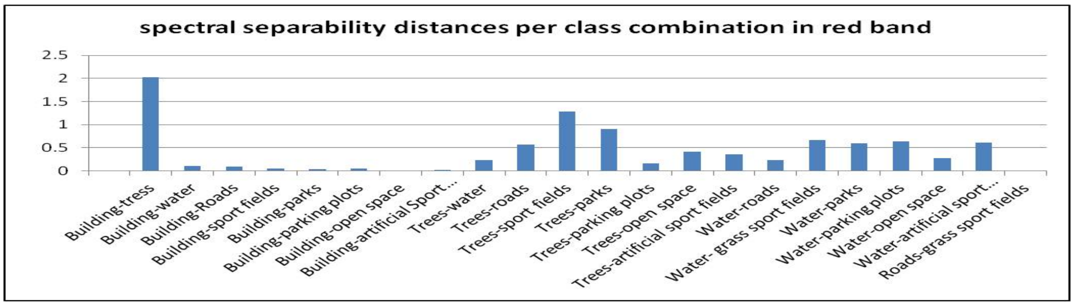

Figure 3.10: The band contains relevant spectral information that enabled the separation between buildings and water, tress and grass sport fields. The remaining classes tend to overlap due to very small spectral distances.

Figure 3.11 shows that the green band contains significant information describing trees and water. However, the separability distances of 0.065 reveals that the separation of building roofs from roads is extremely difficult. The green band contains very few spectral information characterizing parking plots and buildings and the separability distance between the two features is approximately 0.070.

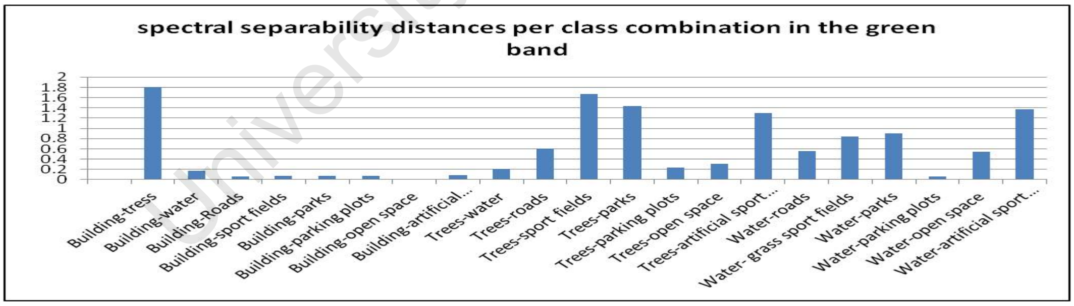

Figure 3.11: The band holds significant spectral information describing trees,grassland and green roofs. Trees and buildings were separated by a quite good spectral distance of 1.0802. Grassland(sport fields) and trees were also separable with a distance of 1.670. Roads and building roofs exhibit similar spectral properties that made the separation difficult as the spectral distance between the two classes was 0.065.

Figure 3.12 shows the different spectral distances in the redness spectral index. The combination enabled the separation between water and roads with a distance of 2.461. The good separation between water and sport fields, recreation parks, open space, artificial sport grounds was clearly illustrated by the respective spectral distances of 5.653, 7.744 , 2.290 and 7.804. The largest separability distance was found between water and parking plots.The two classes can easily be separated at some stage in image classification. In contrast, the combination does not hold significant spectral information characterizing trees and buildings. The two classes are highly confused with other classes in the scene,with separating distances near zero.

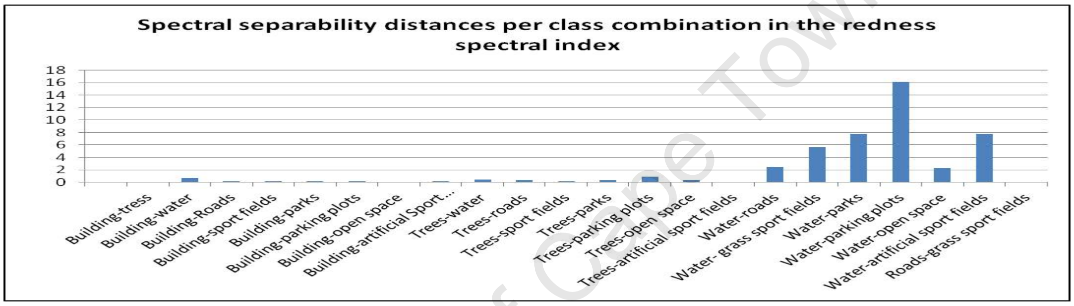

Figure 3.12: Parking plots, grassland, recreation parks and artificial sport fields exhibited high spectral reflectance in the redness index and could be separated from water. Open space and buildings are examples of objects difficult to separate in this spectral index.

Significant spectral information characterizing grassland (sport fields), recreation parks, parking plots and artificial sport fields is enclosed in the greenness band combination as illustrated in figure 3.13. The greenness index allows the separation between sport fields, recreation parks, parking plots, artificial sport fields and water. The band combination holds poor spectral information characterizing building roofs, trees, roads and open space.

The blueness band combination separated recreation parks from trees with a distance of 2.130. The index(figure 3.14) showed a separability of 21.996

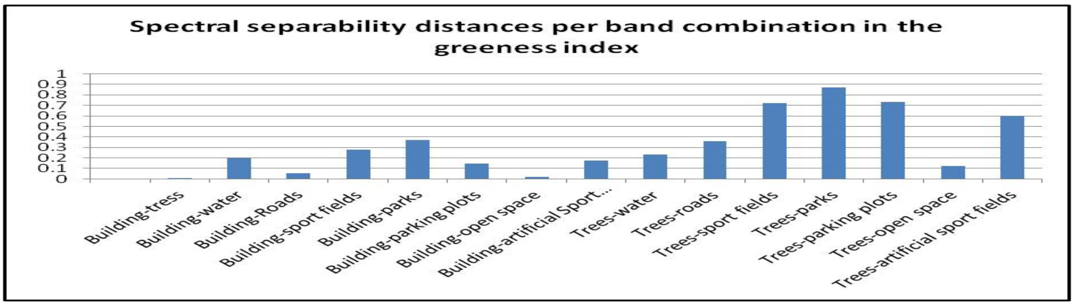

Figure 3.13: Roads and buildings can not be separated in this index with a spectral distance less than 1. Open space cannot be separated from buildings and trees because of the very small spectral distances that separate them. Building and trees exhibit similar spectral properties in this band as shown by the distance close to zero.

between artificial sport fields and water. A spectral distance of 0.64 distinguished trees from water. The separability distance of 20.312 between recreation parks and water was excellent compared to blue, green and red bands

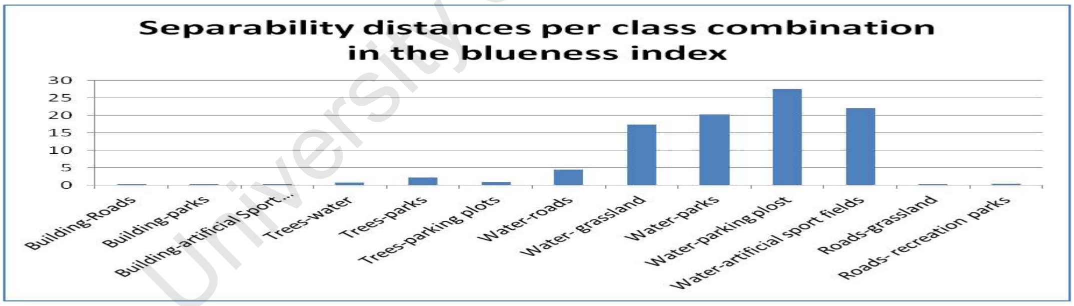

Figure 3.14: Recreation parks can be separated from trees in this band combination. Conversely, building roofs present high spectral similarities with roads, recreation parks and artificial sport fields, making their spectral separation very difficult during classification based on pixel values.

3.3.1 Variation in geometric relations

The results in the figure 3.15 indicate the different mean distances separating buildings within 92 urban sub-scenes composing the study area. Shorter inter-buildings distances characterize high density areas. In contrast, large inter-building distances characterize low density scenes. The urban sub-scene 1 shows a more dense residential area described by a short inter-building distance approximating 13 m. In contrast, the urban sub-scene 4 which has an inter-building distance of about 33 m illustrates a low residential area. Both examples show the significant function of distance relations in describing building patterns.

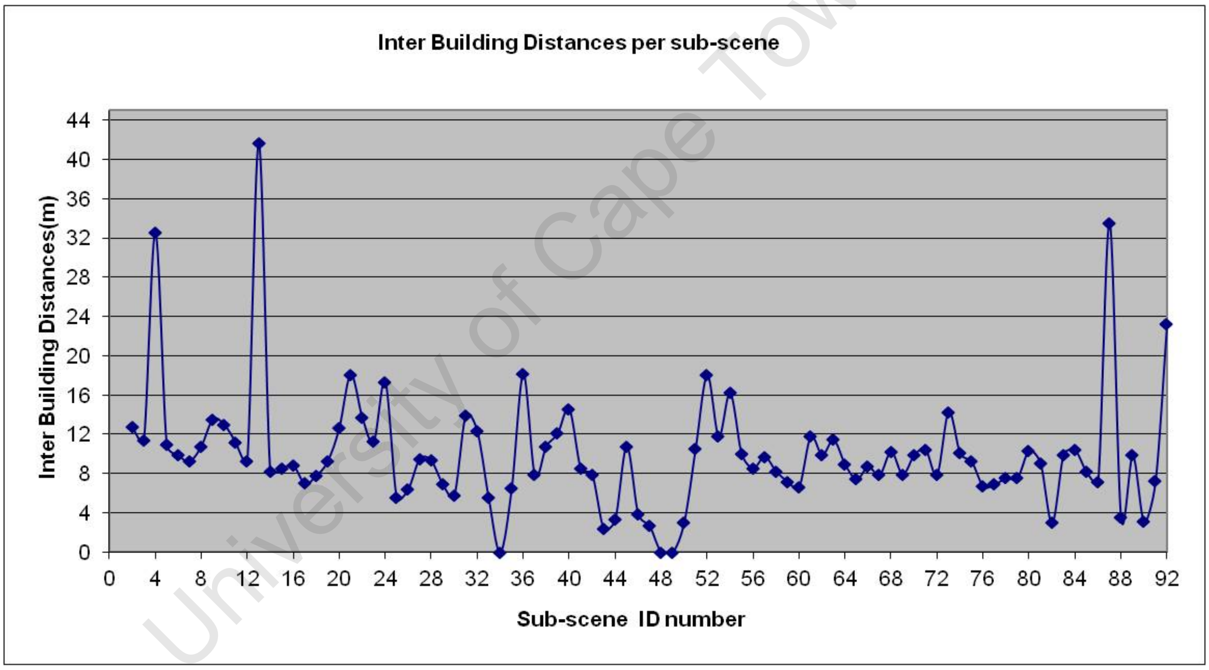

Figure 3.15: Sub-scenes from 14 to 87 are examples of high density areas with number of buildings per area greater than 30. Sub-scenes 4, 13 and 87 are examples of low residential areas with number of buildings per area less than 30.

Figure 3.16 illustrates the different distance relations between roads and buildings in the study area. High residential buildings were identified near primary roads and low residential buildings near secondary roads. Commercial building, used for conducting commerce and other mercantile business,were recognized in centre of towns, near primary roads. In contrast, industrial buildings, described as non-residential buildings of major heavy and light industrial-related infrastructure, were identified in the periphery of urban areas and adjacent to major roads. Formal settlements characterized by formal detached housings and flats, were discovered near formalised street patterns, tarred or gravel roads. Some group of trees were identified along roads and close to individual buildings in residential areas.The distance between buildings and trees varies from 6 to 20 m. Swimming pools were recognized at distances varying between 10 to 25 m from individual buildings in residential areas.

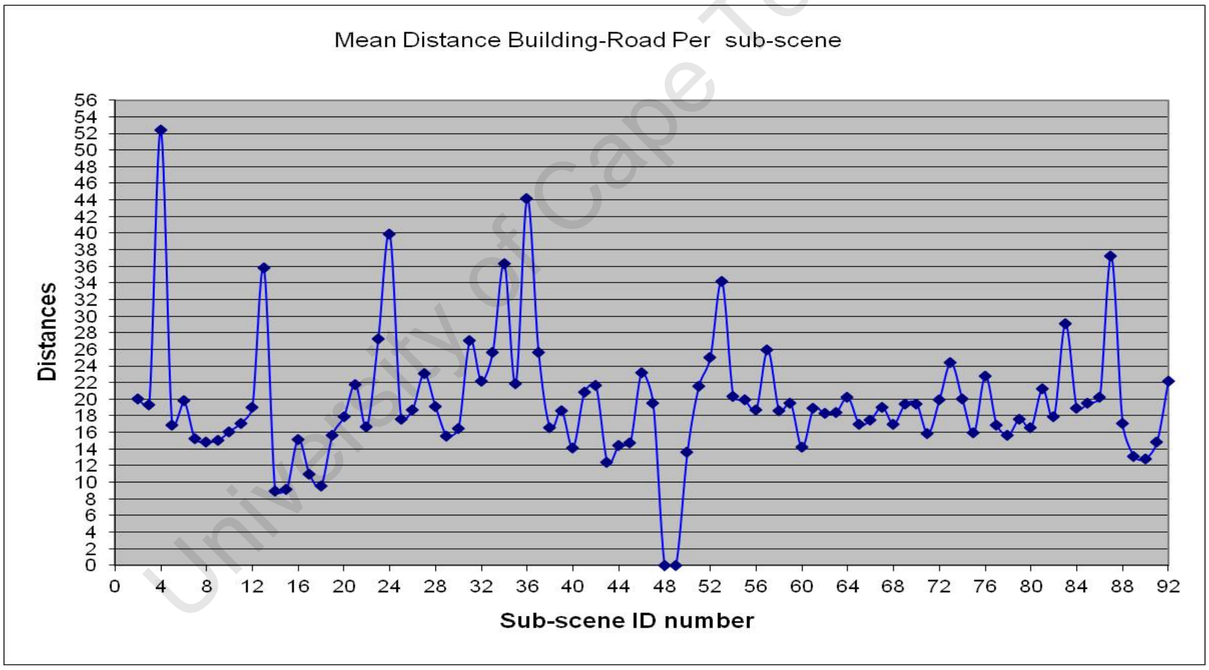

Figure 3.16: Road-building distances: building patterns 4, 12, 24, 34, 36, 52 and 87 are examples of patterns located away from roads, whereas building patterns 16, 18, 28, 40 and 64 are examples of patterns located very close to roads.

The nearest neighbour dispersion index was tested on 92 urban sub-scenes in the study area. The degree of spatial organisation within a scene was assessed based on the scale proposed in figure 3.5. Table 3.3 gives the labelling results of the technique on 11 samples of sub-scenes in the study area. An overall analysis of the findings revealed that 84.09 percent of sub-scenes are composed of regularly arranged housing units. The degrees of clustering found were revealed with a tendency to 2.4. In contrast, a few scenes were found to be composed of clustered and randomly arranged housing units. Only 6.82 percent of scenes investigated were reported to be composed of randomly arranged housing units, whereas 9.09 percent were composed of more clustered housing units.

Table 3.3: Subset of 11 samples of building patterns analysed using the nearest neighbour distribution index : Building pattern numbers 1, 9, 17, 23, 46 and 60 revealed indices closer to zero and were described as composed of clustered buildings. Building patterns numbers 3, 12,20, 39 and 91 indicated indices greater than 1 and were classified as composed of regularly structured building. Building patterns 23 and 60 had indices approximating 1 and were classified as having randomly arranged buildings.

Building patterns	Area in km^2	Building counts	Mean NN distance	S. organization index	organization types
1	40000	40	12.605	0.797	clustered
3	50000	13	32.523	1.048	regular
9	9294.824	9	12.992	0.809	clustered
12	30000	5	41.639	1.075	regular
17	20000	23	7.788	0.528	clustered
20	20000	24	18.112	1.255	regular
23	50000	35	17.371	0.919	random
39	5507.730	13	14.611	1.420	regular
46	6997.105	25	2.682	0.321	clustered
60	10000	15	11.764	0.911	random
91	70000	39	23.219	1.096	regular

The figure 3.17 gives three examples of building patterns analysed with the nearest neighbour technique.

Figure 3.17: Three examples of building spatial arrangement in Cape Town : A: tendency to random spatial organization,B: Tendency to linear clustering organization, and C: Tendency to more regular spatial organization.

Table 3.4: Subset of high level samples derived from different residential scenes. Eight high level features are shown because of the large amount of data as the study was done on 92 residential sub-scenes. Scenes with building counts larger or equal to 30 were classified high density scenes. The letter M stands for mean in the table, MIBD stands for mean interbuild distance, MBRD stands for mean building-road distance and MSV stands for mean spectral value

Building count	M.perimeter	Std Perimeter	M. area	M.compactness	MIBD	MBRD	MSV
40	66.34	18.70	225.23	0.65	12.82	20.01	190.87
39	66.99	16.56	251.00	0.69	11.36	19.28	185.46
13	61.99	24.04	223.53	0.68	32.52	52.42	194.76
30	66.28	17.82	226.48	0.63	10.92	16.87	199.81
20	73.14	10.35	270.79	0.64	9.86	19.79	200.48
18	69.67	8.08	250.33	0.65	9.29	15.19	199.04
15	63.09	11.58	214.87	0.68	10.79	14.83	203.85
49	68.46	21.67	225.85	0.62	13.93	27.00	230.60
23	66.24	21.27	238.20	0.66	12.39	22.15	211.58
31	55.98	28.057	155.93	0.59	5.51	25.63	217.32
35	75.25	81.23	348.01	0.60	18.16	44.18	206.66
10	73.85	31.37	241.79	0.62	12.10	18.61	204.14

Table 3.4 shows a summary of high level features extracted from the different residential scenes. Each residential scene was digitized in order to extract the outlines to measure the area, perimeter, compactness, inter-building distances, building roads distances. The mean spectral values feature was used instead of individual values in red green and blue bands as the deviation between bands was very small. The value extracted, which is representative of the group of buildings within each scene was calculated by dividing the total spectral values by the number of buildings from which the samples were extracted.

The different shape compactness values found show that most of buildings have a regular shape with values ranging from 0.59 to 0.69. Those values were located between 0, characterizing linear objects and the value of 1 characterizing circular objects. This range appears to characterize rectangular objects. Medium size buildings were identified within both high and low density scenes. The shortest inter-building distance was found in a high density scene. Buildings in low density scenes were located at larger distances from roads compared to buildings in high density scenes. More explanations on the findings of this study are discussed in the following section.

3.4 Discussion and conclusion

In this chapter, a model for the description of urban scenes has been presented. The system is made up of: (1) shape and size extraction modules, that adaptively model the shape and size context of different objects within an urban scene, (2) a geometric relations module that models the different distance relations that exist between objects within an urban scene,(3) a spectral module that models the spectral scope of certain objects present within urban scenes and (4) a spatial organization module that models the different spatial organization of objects within urban scenes. As follows, precise hierarchical scene description is established from the pixel level to the global scene scale. Each object is characterized by a vector that includes its spectral characteristics, shape, size, geometric relations and spatial organization, which define a multilevel description of objects. Depending on the level considered, different kinds of features were derived to characterize objects with the most relevant attributes. It is crucial that all features associated with pixel level, object level and scene level modelled in this study, are jointly considered in the classification phase to label urban object. The hierarchical description of the scene from pixel characteristics to pattern organization, allows the capturing and exploitation of the entire information in the scene.

Experimental results obtained in terms of the spectral context analysis confirm the efficiency of the technique used as the findings adhere to those reported by Weng and Quatrochi(2007). The size characteristics results found in this investigation ,adhere to the building size thresholds proposed by Meng et al., (2008). Objects shape models found, particularly for building units, adhere to the conclusions reported by Sirmacek et al., (2010). Distance relations between objects within a scene constitute relevant information in object recognition and discrimination as reported in Agouris et al., (2000). The nearest neighbour module proposed in this study was reported to be a relevant tool for pattern recognition (George et al., 2006). The module enabled the distinguishing of three groups of building units in the study area, namely clustered, randomly distributed and regularly arranged. The module proposed in this study provides advantage over the TOSS technique proposed by Williams & Wentz, (2008). In fact, in addition to object size, proposed to separate different urban patterns, the technique tested in this chapter considered objects spatial arrangement, pixel reflectance and shape signatures, to describe a scene.

Chapter 4

Optimal Scale Parameter selection for a Multi-scale Segmentation.

4.1 Introduction

Multi-scale image segmentation is the partition of an image into spatially continuous, mutually disconnected and homogeneous regions at various segmentation levels (Pekkarinen, 2002). Multi-scale segmentation was initially investigated by Woodcock and Strahler (1987). In the context of image analysis, scale is defined as the level of aggregation and abstraction at which an object can be clearly described (Benz et al, 2004). Multi-scale segmentation starts considering each pixel as an object and merges them to create larger objects based on homogeneity thresholds defined by the analyst (Blaschke, 2010). Multi-scale segmentation provides high level object features, compared to single scale segmentation. In fact, objects generated from this type of segmentation hold additional attributes such as mean value per spectral channel, distances to neighbouring objects, size, as well as shape characteristics. Multi-scale segmentation in e-cognition has the advantage of considering homogeneity criteria such as colour, shape compactness and smoothness, during the creation of image objects. The technique offers the possibility of varying the size of output segments and creates object hierarchy levels that facilitate their accurate extraction. In the absence of objects hierarchy levels, each object will be created from scratch and no topology relationships will be built between objects produced at finer scales and those generated at coarser resolutions. Creating image hierarchy allows each segmentation level, except the first one, to be derived from previous levels. Multi-scale segmentation reduces the spectral variability within scenes and also improves the extraction of objects of various size. However, the quality of objects created from segmentations, relies on the choice of scale parameter, and this choice mostly relies on subjective series of trial and error (Meinel & Neubert, 2004; Kim et al., 2008).The aim of this chapter is to propose a technique that improves the selection of optimal segmentation scales by revealing thresholds that produce segments with low internal and high inter-segment variabilities.

4.2 Material and Methodical Approach

4.2.1 Material

For this investigation, three different subsets of Cape Town urban area, in South Africa, were used to test the technique(Fig 4.1). The areas were extracted from a 0.5m spatial resolution RGB[1] Digital aerial photograph. The choice of these areas was motivated by the fact that each scene is dominated by a different type of urban land covers, ranging from small sized buildings to vegetation extents.

[1]Red Green Blue

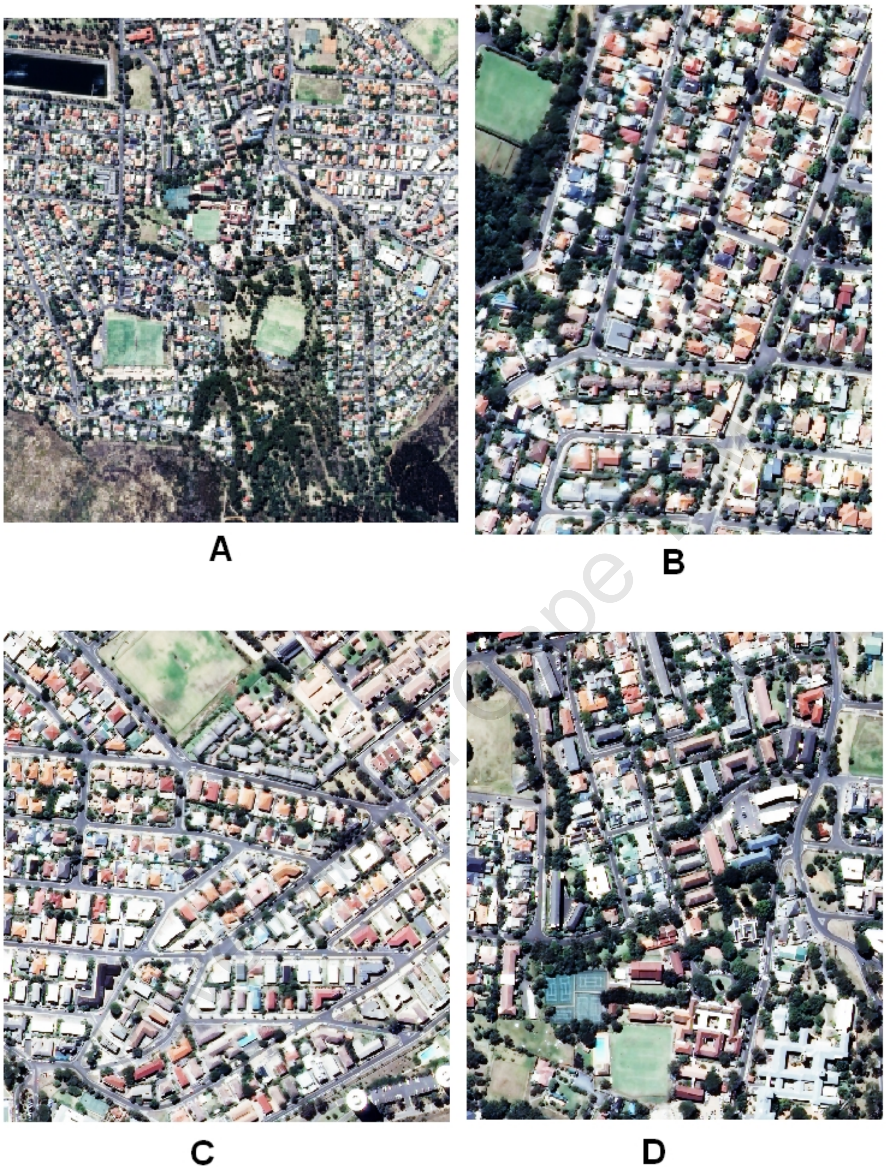

Figure 4.1: A :Original Cape Town urban scene, B: Residential area of Kensington, composed of individual housing units of small and medium size, a large road network, trees and recreation areas. C: Residential area of Vredelhoek. The area also contains a large road network, grass land, artificial sport fields and trees. D: Oranjezicht residential area.

The merit of a segmentation is influenced by the type of land cover investigated and the number of spectral signature considered. This investigation used 1464 distinct spectral signatures samples and 1464 different segment size measures in order to evaluate internal and inter-segment variance and determine the best segmentation parameters. The different urban scenes were beforehand partitioned into several segmentation levels using the multi-resolution algorithm in eCognition. Once segmented, the different objects spectral and size statistics were collected and organized within tables in order to compute segments variances and spatial autocorrelations. The table 4.1 below gives a subset of statistical data extracted from Kensington residential area.

Table 4.1: A subset of the 1464 spectral and size samples used to compute internal segment homogeneity and spatial autocorrelation of Kensington urban area.

Building brightness	Areas	Non building objects brightness	Building brightness	Area	Non building objects brightness
129.95	472	143.43	135.4	434	103.8
147.45	263	14.46	227.97	827	228
209.29	437	14.83	233.63	510	141.75
229.35	520	176.79	197.15	275	80.72
222.11	235	36.14	193.5	300	28.31
148.59	251	229.31	229.54	877	78.59
128.79	180	57.12	209.69	587	32.32
202.14	173	23.61	159.22	523	205.44
238.37	615	152.24	226.01	557	45.67
180.91	117	45.06	200.59	254	17.68
196.61	267	20.83	176.32	261	61.44
166.24	366	56.63	106.67	200	51.96

4.2.2 Segmentation of images

The different segmentations were performed with 13 distinct scale parameters in eCognition, using the multi-resolution algorithm. Multi-resolution algorithm starts with one seed pixel object and merges each pair wisely with neighbouring ones to form larger objects until the homogeneity thresholds defined were reached. These homogeneity thresholds are controlled by the scale parameter which influences the size of output segments. The different image bands which are very relevant in terms of spectral information, were weighted at 1 so that the segmentation algorithm takes into account all the spectral information made available by each band. The remaining parameters including shape, colour, compactness, smoothness were weighted at 0.5 for equal influence on the results. The figure 4.2 shows an overview of the techniques used in this study.

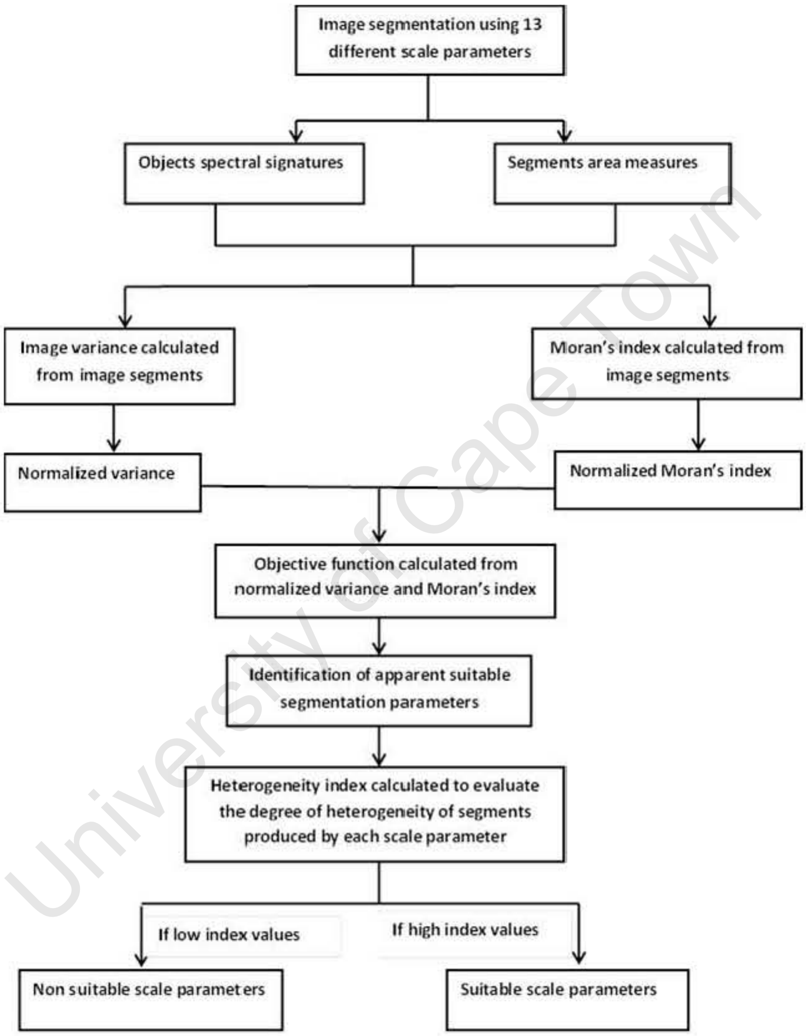

Figure 4.2: Overview of the methods used: Spectral and area measures are extracted from individual segments. These measures are then normalized and used to compute the objective function that identifies potential scale parameters. The potential scale parameters identified are then evaluated using an heterogeneity function.

4.2.3 Identification of suitable scale parameters

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

$$Variance = \frac{\sum\limits_{i=1}^{n} a_i.v_i}{\sum\limits_{i=1}^{n} .a_i} \tag{4.1}$$

In equation (4.1), v_i is the variance and a_i is the area of the segment i. The internal segment variance was weighted with segment areas in order to put more weight on larger segments and limit possible instabilities caused by smaller segments.

The inter-segment variance expressing the spatial autocorrelation between neighbouring segments was computed using Moran's Index. This index was chosen because it is a reliable indicator of statistical separation between spatial objects (Johnson Xie,2011):

$$MI = \frac{n \times \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} W_{ij} \times (y_i - y)(y_j - y)}{(\sum\limits_{i=1}^{n}(y_i - y)^2)(\sum\limits_{i \neq j}\sum)W_{ij})} \tag{4.2}$$

Where $(y_i - y)$ and $(y_j - y)$ are the deviations from the means. W_{ij} is a measure of the spatial proximity of adjacent image objects. In this case the value of 1 was considered for W_{ij} as segments produced by the segmentation are adjacent objects (Espindola et al.,2006). The Moran's Index captures the difference between mean values of each segment and their neighbours. Low values of the index indicate high inter-segment heterogeneity, which is an advantage for image segmentation.

In order to equally consider the internal segment and the Moran's measurements, the indices were rescaled to a similar range between 0 and 1 using the normalization equation below, proposed in Espindola et al,(2006):

$$\frac{X - X_{min}}{X_{max} - X_{min}} \tag{4.3}$$

.

With X, respective values of image variance and the Moran's index to be normalized. The normalized variance and Moran's index of the area were respectively found 0.75345 and 0.26280.

Once the two statistical indicators normalized, they were combined to generate a function called Objective Function given by():

$$F(V, MI) = F(V) + F(MI) \tag{4.4}$$

Where V is the image variance and MI is the Moran's Index. In order to determine the suitable segmentation scale parameters, the different values of objective function were plotted in Microsoft Excel to produce a curve. MS Excel was chosen based on the familiarity and experience with the tool. The lowest values of the curve indicate the suitable segmentation scale parameters.

4.2.4 Assessment of scale parameters

After the identification of suitable segmentation scales, an assessment function was applied to each identified scale in order to measure the quality of internal segment heterogeneity. This function is defined by the following equation proposed in Johnson and Xie (2011):

$$H = \frac{normalizedVar - normalizedMI}{normalizedVar + normalizedMI} \tag{4.5}$$

Lower values of the index reveal high internal segment heterogeneity and higher values indicate more homogeneous segments.

4.3 Experimental Results

4.3.1 Identification of suitable scale parameters

The table 4.2 illustrates the influence of segmentation parameters on the quality of results. In addition to visual analysis of objects delineations, segmentation quality can be evaluated using internal variance and spatial autocorrelation measures. In order to estimate segments internal homogeneity, weighted variances were calculated, taking into account areas of segments in order to minimize the instabilities of small objects on the overall variance of the image. The weighted variances achieved with scale parameters of 15, 52, 35, 55, 65 and 85 revealed that small scale parameters produced objects with low internal variance. This means that neighbouring objects at these scales are still similar and objects have not matched yet their real world corresponding in the image. Similarly, very large scale parameters produced more internally heterogeneous objects,with high spatial autocorrelation between segments, meaning that objects contained more than one category. For instance, high spatial autocorrelation was achieved with the scale parameter of 260 associated with a large value of the Moran's index. The objective function that combined internal and inter segment variances increased with the normalized Moran's index. The highest values of the index were obtained with normalized Moran's index values greater than 0.93. Values of the index have been found greater than 1 with scale parameters of 25, 65, 150 and 260, pointing up high spatial autocorrelations between segments. In contrast, good separation between segments was achieved with scale parameters of 15, 35, 55, 85, 100, 120, 180 and 200.

Segments with low internal variance and low spatial autocorrelation should be expected from optimal segmentations. Low internal variance and spatial autocorrelation(segments statistically different) are characteristics of low values of the objective function. The objective function calculated for the three study sites shown in figure 4.3. reveal the scale parameters of 15, 35,100, 120, 180 and 200 as suitable for segmentation. The curve presented a normal distribution with succession of highest and lowest points. Kim et al.,(2008) described suitable segmentation scale parameters as being associated with negative Moran's index values and the spatial statistics associated with each identified scale were located below zero as illustrated by the table 4.2 above.

Table 4.2: Weighted variances and Moran's indices produced by each investigated scale parameter.

Scale parameters	Weighted variance	Moran's index	Normalized Variance	Normalized Moran's Index	Objective Function
15	1.7932873	-0.07855717	0.77612997	0.34910976	1.12523972
25	1.648730041	1.1792241747	0.73091067	0.964045461	1.69495614
35	1.59351264	-0.33737189	0.753451181	0.262806282	1.016702283
55	1.674530841	0.026208054	0.67873569	0.94206933	1.6208053
65	1.70178086	1.042506743	0.74495658	0.943739052	1.688695631
85	1.67215105	0.359926887	0.54461091	0.867445256	1.4120562
100	1.795287344	-0.12242074	0.90272707	0.040614401	0.94334108
120	1.792175689	-1.15040280	0.62447331	0.09476837	0.71924168
150	1.67331783	1.176001299	0.63966041	0.99267777	1.63233812
180	1.654452769	-0.03452559	0.69525634	0.553309844	1.248566183
200	21.3007068	-0.00620757	0.91439718	0.34392102	1.258318202
260	1410.920807	18.32538678	0.43182987	0.54792159	0.97975466

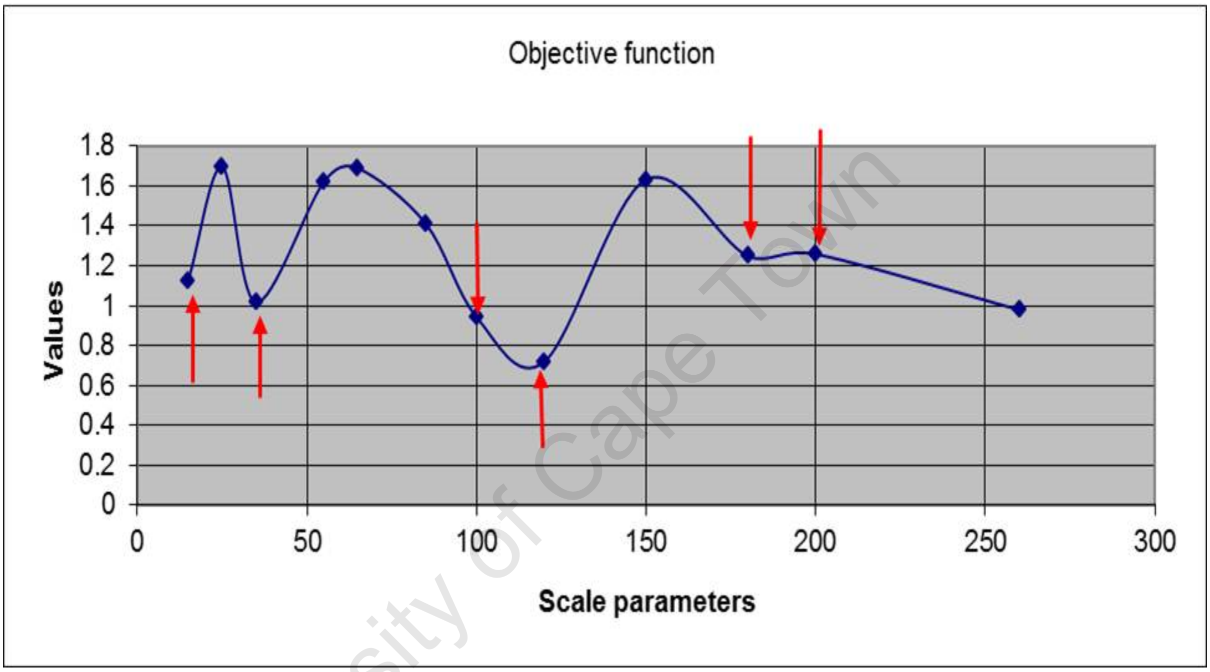

Figure 4.3: Objective function identified seven low values suspected to be associated with suitable scale parameters but only six of them are related to negative Moran's index, characteristics of suitable scales (table 5.2). Scales of 15, 35 and 100 were found appropriate to segment Kensington urban scene, mostly dominated by small and medium size buildings; The scale parameter of 120 was found suitable to segment Vredelhoek area dominated by a large road network and the scales of 180 and 200 were found appropriate to segment forest extent in Oranjezicht area.

4.3.2 Refining under and over segmentation

In order to assess the different identified scales, an heterogeneity function was applied, in order to evaluate the degree of heterogeneity within different segments produced at different segmentation levels. The normalized heterogeneity index gives values closer to 1 when describing more homogeneous image objects. In contrast, values closer to zero characterize objects with high internal segment heterogeneity. The figure 4.4 presents the different heterogeneity scores associated with each scale.

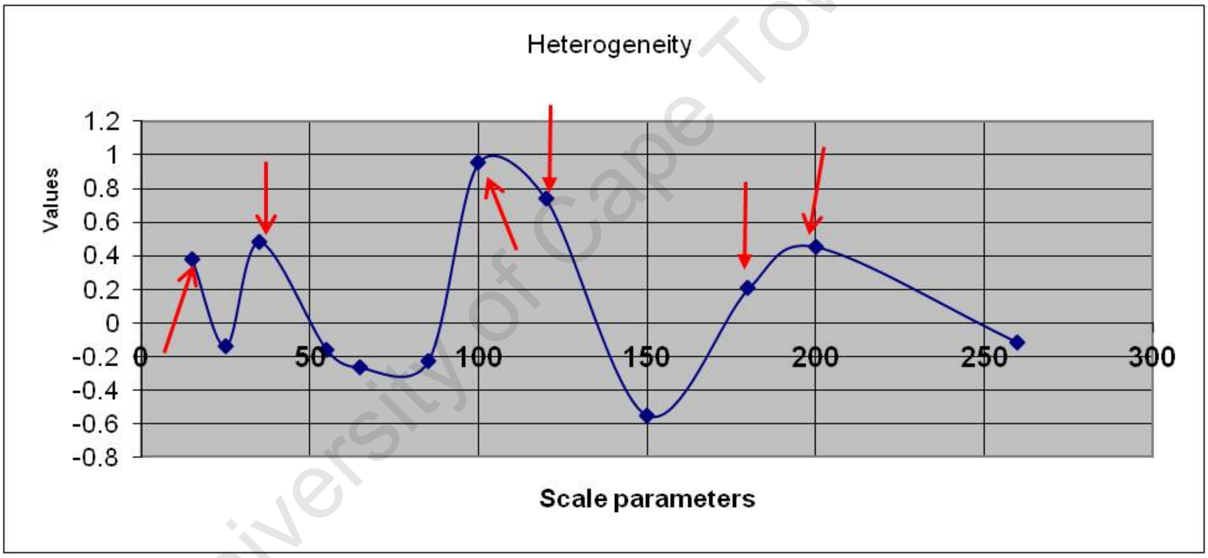

Figure 4.4: Heterogeneity function showing segments internal homogeneity at different segmentation scales. The higher the value of the index, the more homogeneous are the segments produced. Scales of 100 and 120 produced better segment homogeneity with higher values of the index. Scales of 35, 180 and 200 produced good enough segments homogeneity whereas scales of 25, 50,55, 65, 150, 260 produced poor results with high internal heterogeneity.

The figure 4.5 illustrates visual segmentation results of the area of Kensington. The scale parameters of 35 and 100 successfully delineated small and medium size buildings. Most of objects at these scales approximated their real world equivalent. The scale of 25, 50, 55 and 65 over-segmented the image, segments delineations were out of the real world object boundaries.

Figure 4.5: A: original Kensington scene, B: segmentation at scale of 10 showing building roofs partitioned into several smaller segments, C:segmentation at scale of 55 showing over and under- segmentation of the areas and D: segmentation at scale of 100 showing building roofs segments approximating their real world corresponding.

The segmentation performed on the Oranjezicht scene produced a satisfactory delineation of the road network, when performed with the scale parameter of 120. At lower scales, the process produced small individual road segments difficult to classify at a later stage. Figure 4.6 shows the visual results of roads segmentation of the study site.

Figure 4.6: A: original Oranjezicht scene, B: segmentation at scale of 25 showing roads segmented into several smaller objects, C:segmentation at scale of 35 still showing small segments of roads and D: segmentation at scale of 120 showing road segments approximating their real world lengths.

In figure 4.7, the scale parameter of 25 produced individual trees with the size approximating some small buildings, complicating their separation from other objects. Forest stands, easy to extract based on size and shape,were obtained at scale of 200 over the area of Vredelhoek. Larger scales tested produced mixed objects composed of more than one class.

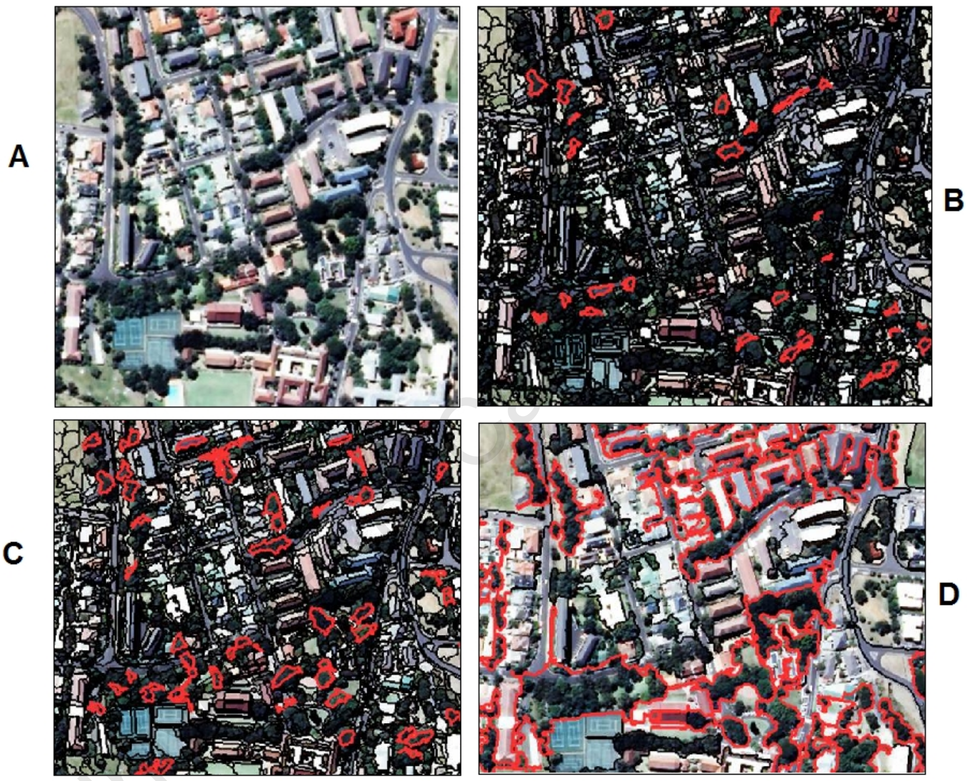

Figure 4.7: A : the original Vredelhoek scene, B: segmentation results at scale of 25 showing individual trees, B: segmentation at scale of 55 still showing small trees segments and D: segmentation results at scale parameter of 120, individual trees have been grouped to form larger forest extents, suitable for classification.

When comparing the visual results of segmentations done at scale of 120 and 150, it was revealed that segments produced at scale of 150 contained more than one object types as illustrated in figure 4.8.

A  B

Figure 4.8: A: segmentation results performed with the scale parameter of 150, the resulting objects contained more than one class including buildings, roads, open space and trees. B: segmentation done at scale parameter of 120, the segment comprises a portion of road that is optimally captured compared to results in A.

4.4 Discussion and conclusion

While the quality of segmentation results relies on the choice of segmentation parameters, the choice of suitable scale remains a challenge and mostly results from subjective trial and error process. Consequently, techniques that enable the identification of appropriate scale thresholds are needed. Introduced by Espindola et al,(2006), internal segment variance and spatial autocorrelation concepts were used to calculate the objective function in order to identify suitable scale thresholds to segment the different scenes under investigation in this study. Instead of using a standard object size as previous researches(Espindola et al.,2006; Kim et al.,2008; Martha et al.,2011; Johnson & Xie,2011), this study considered individual size of segments produced by each scale parameter.

The tool proposed in this study was tested on three subset areas of Cape Town urban scene and the results obtained from the analysis of the objective function agreed with findings of Kim et al,(2008) that suggested that suitable scale parameters should be associated with negative values of Moran's index. The curve of the objective function obtained enabled the identification of six scale parameters suitable to segment the area of Cape Town. These scales were then assessed using the heterogeneity function proposed in Johnson and Xie(2011). The assessment confirmed the homogeneity of different segments produced at the identified segmentation scales. The different objects obtained with the identified scales approximated the size of their real world corresponding. The multi-scale aspect of the approach gives the advantage of targeting different sizes characterising each object within the scene in order to find their appropriate scale of extraction. This is very relevant for multi-scale analysis of high resolution imagery as proposed in recent investigations(Espindola et al.,2006; Dragut et al.,2010; Kim et al.,2008; Johnson & Xie,2011; Martha et al., 2011).

The results found in this study were similar to those of Johnson and Xie (2011), Espindola et al., (2006), which reported a decrease of objective function values as objects tend to reach their real world corresponding within the image. Conversely, these results were different from those of Gao et al.,(2006), who found an increase in objective function values at optimal scales. Their results may have been affected by the spatial structure of their study area which was an homogeneous forested area that generally has high spectral similarity between different vegetation classes. This research was based on a more heterogeneous urban area of Cape Town, which exhibited a high variety in spectral characteristics.

An analysis of the different objective function values showed that segmentation at scale parameter of 120 produced the lowest index value as illustrated in figure 4.3. The Moran's index associated with this scale parameter was very low, reaching the value of -1.15040 as showed in table 4.2. This result revealed that image segmentation done with this scale-produced segments with low internal variance and low spatial autocorrelation. In order to assess the different scale parameters identified from the objective function, a heterogeneity index was applied to each of these scales and the results in figure 4.4 confirmed the suitability of the scale of 120.

Spectral information used in this investigation played a major role in the success of different segmentations. However, it could be an advantage to incorporate some vector data such as cadastral boundaries or building vector footprints, into the segmentation process. In a recent investigation, Hamit Kok, (2005) tested a segmentation of an agricultural area supported by the vector area boundaries and the results were promising.

Besides object spectral properties, object size enabled the author to identify suitable scale parameters. The consideration of segment size in the calculations of object variances gave more weight to larger objects to influence the results, minimizing instabilities of small segments that generally create over-segmentation and produce inaccurate results. This study only investigated 13 scale parameters over the urban area of cape Town, in South Africa. However, more researches are suggested to test the sensitivity of the proposed technique with more ranges of scales and on other land cover types.

This study proposed a technique of scale parameter selection for an optimal segmentation. The approach enabled to objectively compute object internal variance and spatial autocorrelation index. The study used all the spectral bands suggested by the investigation done in chapter 3. Nevertheless, for future investigation, the use of infra-red band would be suggested because of its capability to discriminate buildings from their backgrounds (Small, 2003).

The results of this study revealed that selection of scale parameter based on the objective function could increase classification accuracy as good segmentation leads to good classification. The use of heterogeneity function enabled the author to identify which scale threshold could produced under and over-segmentation results (figure 4.8). By using this combination of techniques, the study overcame the limitations of traditional methods of segmentation scale selection relying on subjective series of trial and errors, which do not always guarantee the success of the process. Associating non-spectral data such as vector building footprint can help to accurately extract them from the segmentation and the investigation of more range of scales would also have improve this results. The following chapter of this thesis will test the scale parameters optimal selection technique on a larger urban scene of Cape Town, South Africa, containing the same land covers in the purpose of a multi-scale context based classification.

Chapter 5

Multi level Image Classification

5.1 Introduction

Recent researches have pointed out that object recognition, which appears to be a natural process for human vision, remains very challenging for computers. This limitation for computers to accurately recognize objects within a scene hails from the fact that object identification is not a local process but rather a global one. The simple operation of identifying a building within a forested area can be very limited if only based on pixel attributes but the inclusion into the process of information such as shape, size and the consideration of adjacent objects such as roads, can improve identification accuracy. In the last decade, traditional classification systems have attempted to recognize objects based on low level pixel information, but results have reported topological errors such as finding building on top of trees or roads within water bodies due to spectral similarities. Associating contextual information into the classification process can prevent such errors and improve the recognition results. Moreover, several types of contexts have been studied in literatures and among them are the following: (1) the semantic context which defines the probability of an object to be found in some scenes and not in others, (2) the spatial context which determines the chances of an object to be in certain positions related to other objects within a scene and not others. Distance relations which illustrate such type of context, can be divided into two main groups : relations that express fixed measures of scale in pixel count or other metric units, and relations providing relative measures such as close to, near to, (3) the scale context focuses on the fact that certain objects types have limited range of size characteristics in respect to other surrounding objects, (4) spectral context which exploits the fact that certain types of objects exhibit specific spectral properties in certain bands, compared to others and finally (5) topological context describes the relationship between an object and its surroundings. Such relations are defined by considering interior and exterior boundaries of an object and membership relations, intersection relations such as touches, overlaps, are illustrations.

This chapter presents a multi-level classification system that takes into consideration the different object contexts available within the scene. In order to assign the different objects to their appropriate classes, the proposed technique was built on the Nearest Neighbour and Fuzzy logic rules. The procedure proposed intended to test the knowledge-based system on aerial photographs, improve classification accuracy of aerial photographs and attempt to handle the difficulties related to spectral similarities in urban areas.

5.2 Material and Methods

5.2.1 The data

The image used for this investigation was an aerial photograph of Cape Town, taken on summer 2009 in order to minimize the effect of cloud cover on the quality of the image as generally observed on images taken during the winter period. The spatial resolution of the image was 0.5 meters which was suitable for urban areas mapping. The true colour image was composed of red, green and blue bands, located within the visible spectrum of the light. Pre-processing operations were already performed on the image before it was collected from the Geomatics department of University of Cape Town. Object spectral characteristics were previously extracted using e-cognition software in order to perform all the spectral analysis. Size, shape and distance measures were extracted from digitized polygons in ArcGIS10 version made available by the Geomatics Department, University of Cape Town. Topology relations were directly derived from image objects produced at different segmentation levels using the multi-resolution algorithm in e-cognition.

5.3 Image segmentation

Semantic knowledge useful for image classification is not always available from raster images. Image segmentation offers the potential of extracting such information by making available object's shape, size and mutual relations characterizing the different objects generated. Given that raster images are composed of pixels, segmentation algorithms group homogeneous pixels pair wisely to form larger object of interest, based on homogeneity criteria defined by the analyst. Among the different homogeneity criteria involved in image segmentation are the scale parameter, the shape, colour, smoothness and compactness factors. The different weight values attributed to any of these values influence the segmentation output. The consideration of different spectral bands made available by the image can also be controlled by attributing a score ranging from zero to one to a given band. The higher is the score, the more influential the spectral band will be. The scale parameter, which is a very relevant factor for image segmentation controls the size of output objects. The higher is the value of scale parameter, the larger will be the objects generated. Since single segmentation produces very limited information regarding the produced objects due to the diversity of object size present in most of complex scenes such as urban areas, multi-scale segmentation techniques were proposed in recent studies(Bruzzone and Carlin, 2006; Lu et al. ,2010). Multi-scale segmentation produces several hierarchy levels of objects as well as mutual relationships existing in the real world. Each resulting object knows its neighbours, not only at a given level but also at finer or coarser levels. This advantage enables to derive object different contexts useful for image classification.

The selection of suitable segmentation parameters for an optimal classification is very challenging and generally relies on subjective series of trial and errors, which do not always guaranty optimal results. A method of selection of optimal scale parameter was proposed in the previous chapter 4. The optimization of scale parameter selection studied in chapter 4 aimed at minimizing the under and over-segmentation in order to increase the efficiency and accuracy of the proposed classification approach. Given the spectral and size variability within urban areas, single scale segmentation techniques will never generate meaningful objects of different size (Martha et al., 2011). Therefore, this study addressed a technique in which the overall analysis of the scene contains several segmentation levels. The potential of e-cognition multi-resolution segmentation algorithm and the lowest values of the objective function were used to segment the aerial photograph of Cape Town, South Africa. Each lowest value of the objective function corresponds to the minimum intra and inter-segment heterogeneity; therefore minimizing under and over-segmentations. The lowest values of the objective function curve are distinct from each other and segments obtained with scale parameters corresponding to these points, approximate their real world correspondent on the image. Segments generated by the scale parameter of 100 were used to begin the classification as this scale delineated better, object contours and produced the best level 1 features. This scale parameter had the potential of outlining small and medium size objects, in comparison to scales of 15 and 35. The scale parameter of 120 was used to create level 2 classification in order to detect large size objects, as well as elongated features such as roads. The scale parameter of 180 was used to generate level 3 classification and gave the possibility to detect tree plantations, parking plots, and grassland.

5.4 Multilevel contextual classification

Bruzzone and Carlin (2006) demonstrated that adding local context into a classification process could improve the results. The figure 5.1 shows the same object represented in two different contexts.

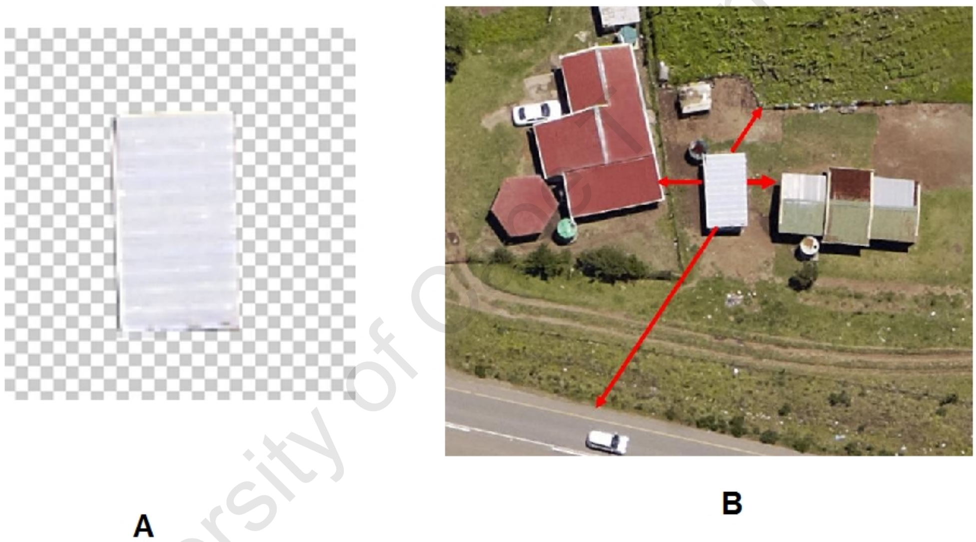

Figure 5.1: A: the type of the object cannot easily be determined due to insufficient contextual information. B: put into its real context, the object was identified as a building. This illustrates the difficulty encountered by computers to recognize unknown objects without context.

In order to model the different spatial contexts generated from the segmentation process, it is important to link the different objects at their scale of representation as well as at inter-scale level within a semantic network. The figure 5.2 below shows an example of a semantic network for the class urban.

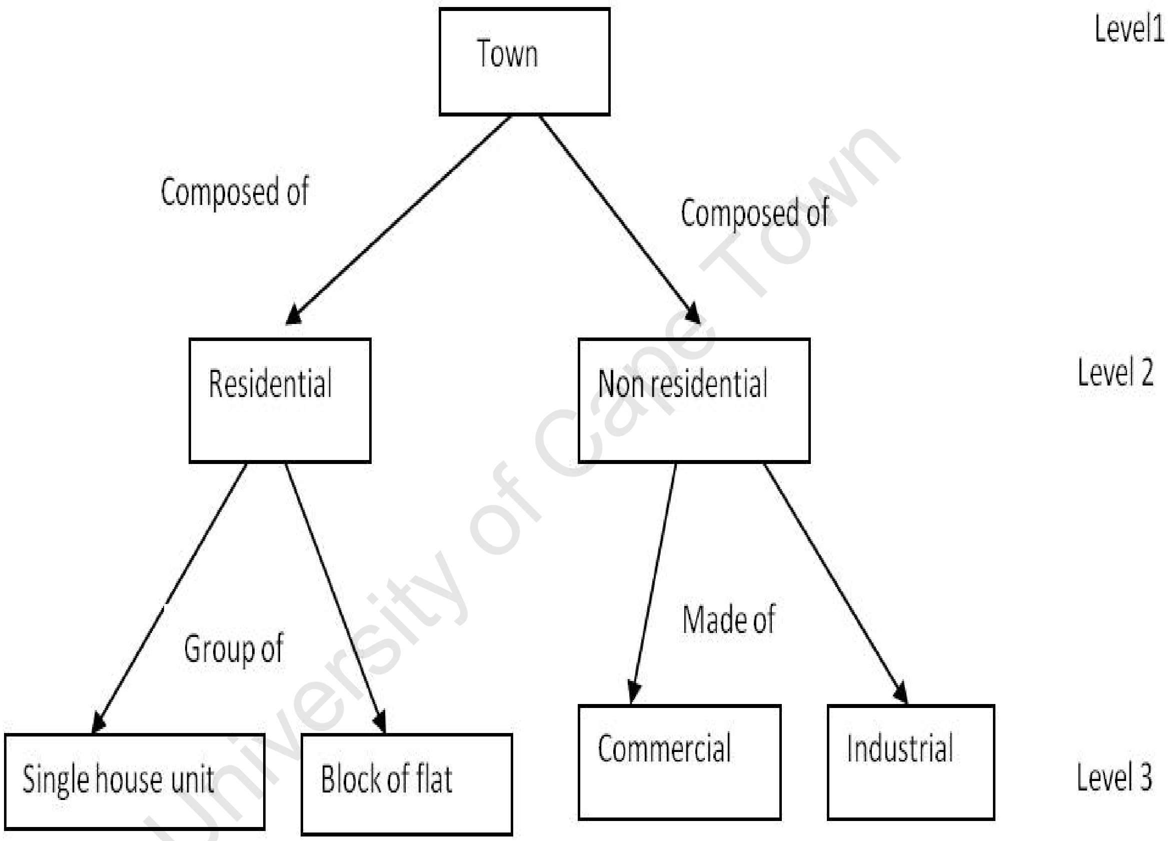

Figure 5.2: An example of a semantic network for the class building. At level one, buildings are observed as a town, at level two, buildings can be observed as residential and non-residential structures. At level three, buildings are perceived as single houses, block of flats, industrial and commercial structures.

Three levels of classification were performed on the high-resolution aerial photograph of Cape Town. Low level objects attributes including size, shape and distance relations produced by the first level segmentation, were mainly used at level one classification. The use of high-level contextual information could only be applied at higher-levels of classification once objects have been labelled. Object based classification approach tries to solve the problem of spectral similarities between different classes. Segments generated from the segmentation contain pixel reflectance values, shape, size and topology, which were integrated into the classification system. In this study area, impervious surfaces are extremely complex to separate as shown in the figure 5.7. Different impervious surface such as building roofs, roads, parking lots, bare ground have different spectral signatures and are confused with other land covers such as bare ground, water and rocks due to spectral similarities. Therefore, proper selection of object attributes for object-based classification is critical. Since impervious surfaces have large spectral variation, no single class can represent all the impervious surfaces. Thus, different impervious surface spectral signatures were selected for classification, representing low, medium and high spectral value impervious surfaces.

The remaining land covers included trees, grassland, open space, shrubs and water. At least 180 spectral signatures were selected to represent the different impervious surfaces. 20 spectral signatures were used to characterize trees, 20 spectral signatures were used to characterize water, 15 spectral signatures were used to characterize shrubs class, and 10 signatures were used to characterize open space. In addition to spectral signature, three shape variables were used mainly shape compactness, length, width/length ratio and one size feature was used mainly the area measure. Contextual information was also used in this research, involving mean distances between objects, membership to a class, bordering relationship. The multi-level classification system used in this study was based on the nearest neighbour rule, because of the flexibility and for its classification accuracy. Three levels of classification were built in order to extract the different land covers. The level 1 or pixel level grouped the different classes into corresponding categories and six categories were extracted using e-cognition software. Here only pixel values, size and shape attributes were used. The level 2 classification segmented the original level 1 classes into nine classes based on contextual information between objects. At this level, built up class was segmented into residential and non residential categories based on size attributes, artificial sport fields were separated from impervious surface group to form a single group based on shape compactness approximating 0.56. Recreation areas were separated from grassland group based on shape and size attributes and their spatial distance in relation to roads, which approximated 20 meters. The level 3 classification used more contexts to build 12 land use classes from level classification. Topological relations between buildings and parking lots enabled the separation between commercial buildings and non-commercial structures. The size of parking lot also played a role in the separation, as commercial structures are generally located near large parking lots. The distance in relation to roads enabled the distinguishing of industrial structure and non-industrial structures. As demonstrated in chapter 3, industrial structures are located in peripheries of town and linked to secondary roads, whereas commercial structures are generally linked to main roads. The following decision tree in figure 5.4 summarizes the different rules used to create the last level of classification. The motive of only presenting a decision tree for this level was motivated by the high classification accuracy produced with the value at 91 percent.

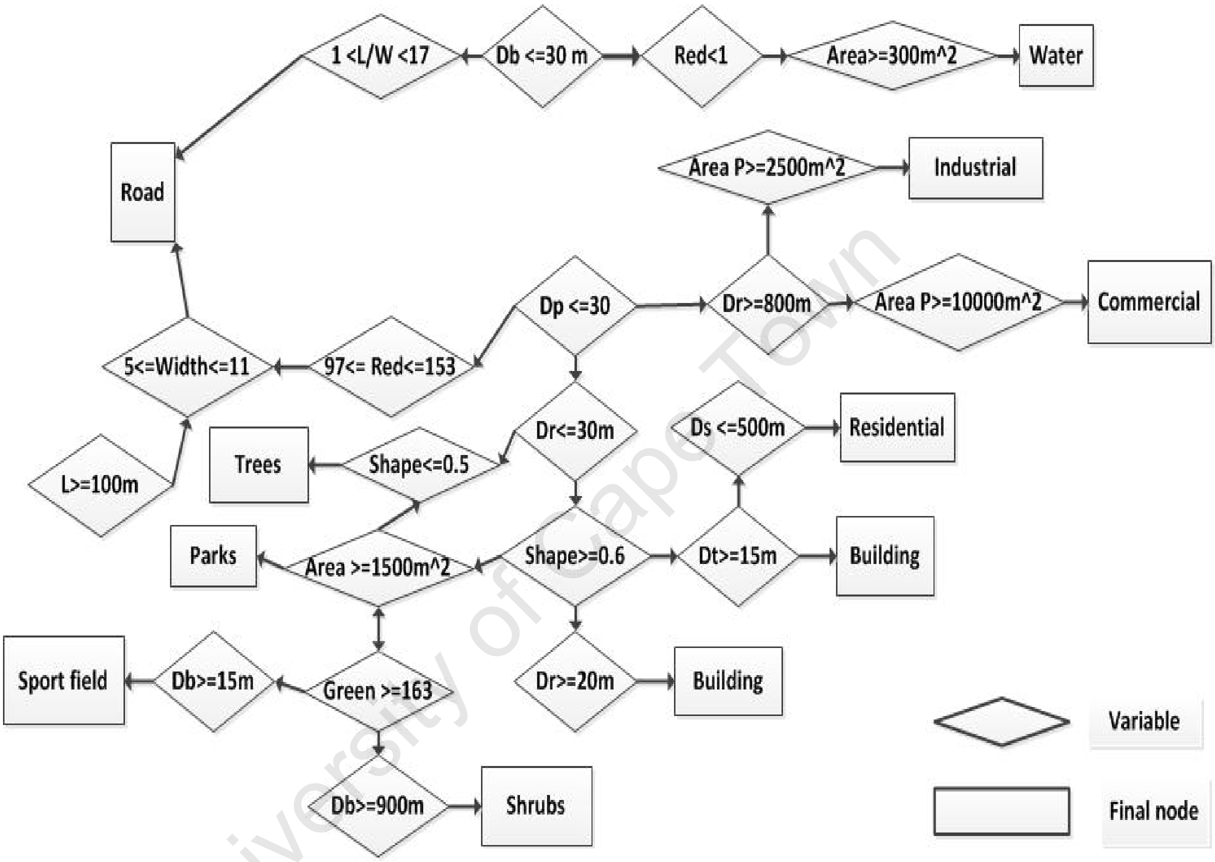

Figure 5.3: Level 3 classification decision tree. Variables and corresponding thresholds are represented by the quadrilaterals. The different terminal nodes are represented by rectangles.The arrows indicate the direction of the decision which are represented by terminal nodes. As example an unknown object located at a distance smaller or equal to 30 meters from any building, with a length over width ratio ranging between 1 and 17 is likely to be a road segment. For a definition of the different variables see table 5.1 below.

Table 5.1: Basic variable used in level three classification decision tree. The type column defines if the variable is spatial or spectral and the layer indicate the band from which the variable was extracted in case of spectral type

Variables	Definition	Type	Band
L/W	Ratio of length over width	Spatial	None
Db	Distance from the object to the closest building	spatial	None
Red	Image object mean reflectance value in red band	Spectral	Red
Area	Area measure of an object	Spatial	None
Dp	Closest distance to parking	Spatial	None
Dr	Distance from the object to road	Spatial	None
Area P	Parking area measure	Spatial	None
Shape	Object shape compactness	Spatial	None
L	Length	Spatial	None
Width	Object width measure	Spatial	None
Ds	Distance to sport fields	Spatial	None
Dt	Distance to the closest trees	Spatial	None
Green	Object mean value in the green band	Spectral	Green

Since impervious surfaces exhibit high spectral variation, no single class could represent all the impervious surfaces. As a consequence, a three level classification was performed on the high resolution aerial photograph of Cape Town. The first classification level consists of eight land cover classes namely Buildings, Water, Trees, Roads, Grass, Artificial asphalt sport fields, Parking lots and open space. At this level, only spectral, shape and size as well as distance measures were useful to identify the land covers from the unknown image objects. The use of high level context information could only be applied at higher level classification once objects have been attributed labels. Impervious surfaces were difficult to extract based only on pixel information due to spectral confusion between features such as building roofs, roads, parking lots, and bare ground, water and rocks. Building roofs for example were identified from the unknown segments using the spectral mean value in red band varying between 100 and 243 and area size ranging between 100 and 300 square meters. Shape compactness was an advantage to support the assessment of building recognition with the range varying between 0.61 and 0.63. Tree plantations were easily isolated in the red band between the values of 15 and 83.8 and in the green band between 23 and 94. Background shadows produced irregular shapes around tree patterns, which were isolated based on a shape compactness values between 0.69 and 0.73. Grassland was firstly isolated based on size which was greater that 500 square meters, as the class exhibits spectral similarity with open space, shrubs and water. In addition to spectral signature, three shape variables were used, mainly shape compactness, length, with/length ratio and one size feature was used mainly the area measure. Roads were extracted based on a $\frac{Width}{Length}$ ratio smaller than 1.6 and a length measure greater than 100 metres. Artificial asphalt sport fields were distinguished from the unknown segments based on their high reflectance in the blue light channel and their high shape compactness signature tending towards rectangular shapes. Parking plots were isolated from the unknown objects based on their high reflectance in the green light channel. Their extraction was facilitated by the previous isolation of roads, artificial asphalt sport fields and building roofs which share similar spectral characteristics and sometimes shape signature.Water bodies were isolated from the unknown objects based on their low spectral reflectance in the red light channel. And finally, the remaining unknown image objects were classified as open space class.

At a second classification level, the attention was mainly on the separating non residential buildings from the building class. The area size measure was used to isolate large buildings likely to be non residential buildings from residential buildings of small and medium size. In fact, non residential buildings were characterized by an area size greater than 300 square meters, whereas, residential buildings were described with an area size varying between 100 and 300 square meters. In order to minimize the confusion between a few large residential buildings and non residential buildings, shape compactness measures were used as non residential buildings are elongated buildings with low shape compactness values. All other land cover classes were maintained for separation in the third classification level.

The third classification level aimed at separating individual land covers into land use classes, based on contextual information. For instance, commercial and industrial buildings were separated from the non residential building group based on their distance relation to residential buildings and parking lots. Moreover, the global context of both object types was also helpful to assess their extraction as commercial buildings tend to occur in the centre of town whereas industrial buildings tend to occur near the periphery of town. Both classes were located at a mean distance less than 20 meters from parking lots. The size of parking has played a major role in discriminating between the two classes as commercial buildings are generally located near large parking lots with size greater than 800 square meters whereas industrial buildings were located near medium size parking lots. The proximity to main roads was also relevant to assess the separation between the two features as commercial buildings are generally connected to main roads, whereas industrial buildings are connected to secondary roads. The global context of the scene was also used to assess the classification as industrial structures generally occurred at the periphery of town whereas commercial buildings are near the city centre. Artificial sport fields were separated from parking lots based on the " borders to" topological relation. In fact, parking lots have boarders more connected to roads than artificial sport fields. Educational buildings were separated from the non residential class based on the proximity to artificial sport fields, as most of educational structures are located near sport facilities. To assess the misclassification of parking lots class as artificial sport field class, the membership relation was used. All objects classified as artificial sport fields which did not shape any borders with main roads were attributed a membership weight of 0 to reclassify them into the appropriate category. Open space segments were separated from buildings based on the shape compactness. Open space objects were characterized by irregular shapes whereas buildings had a more regular shape description. The "boarder to" trees enabled to separate open space objects from building segments despite spectral similarity. Open space segments were likely to share borders with trees than buildings. Swimming pools and dams were discriminated within the water class based on their respective size and global context within the image as swimming pools tend to have smaller size and located closer to residential buildings and dam tend to occur at the periphery of town around agricultural areas and have larger area sizes. Recreation areas were separated from sport grassland within the grassland group based on their shape compactness as recreation parks have more irregular shape characterized by lower shape compactness values than sport grass land. Contextual information was not helpful to discriminate between the two sub-classes as they tend to occur in similar context within urban areas.

5.5 Experimental results

For all subset images of Cape Town, scale parameters corresponding to the lowest points of the objective function curve were selected to perform segmentations. A visual assessment of the three segmentation levels is presented in figure 5.4. The technique accurately identified appropriate scale parameters as objects of small and medium, large size buildings, road network; trees and water bodies have been accurately delineated. For the level 1 segmentation, the scale of 100 was selected because this scale correctly delineated individual buildings and block of buildings. However, scale parameters of 15 and 35 produced less meaningful segments than the scale of 100. Comparison of the different visual segmentation results against each other revealed that objects boundaries accurately delineated at lower levels are preserved at higher levels.

Figure 5.4: A: segmentation results at scale parameter of 100 showing good delineation of small and medium size building units, B: image segmented at scale parameter of 120 showing good delineation of large size objects such as large buildings, grassland and parking lots ,C and D: image segmentation results at scale parameter of 180 showing good extraction of road network as well as group of trees.

The experimental results obtained at different classification levels confirmed the advantages of integrating shape, size and contextual information into the classification compared to traditional techniques relying only on pixel information. These advantages were revealed from the good recognition of classes such as trees, grass land, artificial asphalt sport fields. Even though the entire road network was identified at all levels, their spatial continuity in the borders was distorted by the presence of trees along the feature.

In the classification results at level 1, based mostly on pixel information, major land cover misclassification errors involved spectral confusion between buildings and open space, open space and trees, water class and some building shadows. The figure 5.5 shows the complexity to separate impervious surfaces from other classes that share similar spectral properties.

Figure 5.5: Classification results at level 1 on the right hand side compared to the original image on the left hand side. Land covers were extracted based on pixel reflectance,size,shape and distance measures.The yellow circle shows a grassland accurately extracted based on pixel reflectance,size measure and shape compactness.

The use of classes spatial characteristics such as size and shape enabled the improvement of level 2 classifications, as visually illustrated in figure 5.6.

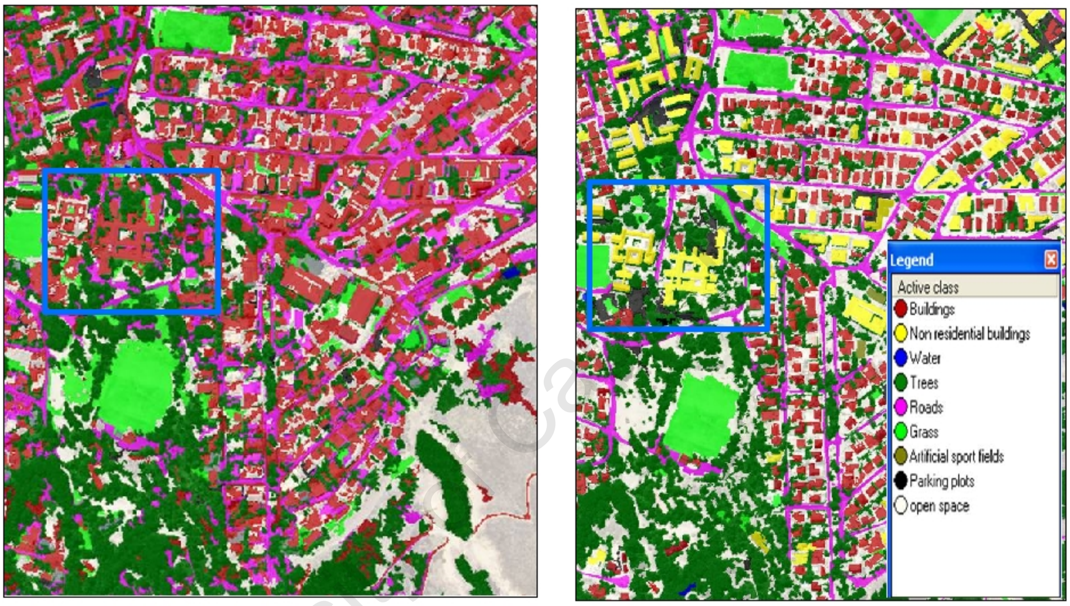

Figure 5.6: The classification at level 2, on the right hand side was improved by the use of context information that enabled the separation of residential from non residential buildings. The distance relation between buildings in classification level 1 on the left hand size and parking plots played an important role in distinguishing non residential from the building class as illustrated in the level 2 classification results.

Level 3 classification showed minor confusions between classes and this advantage originated from the integration of contextual information (figure 5.7). The minor confusion identified occurred between commercial, educational buildings and open space. Artificial sport fields were also confused with open spaces in some areas and the confusion could not be avoided due to poor context of open space as there was a lack of other objects to exploit the spatial relationships.

Figure 5.7: Great improvements can be observed on the image on the right produced at level 3 classification, compared to the image on the left hand side produced at level 2. This improvement was possible by the use of relationships such as membership, border to, super object of,sub object of, used in this level. The red circles show classification improvements of a group of trees using the membership relation.

The following error matrix was built from the final level 3 classification results, in order to measure the degree of correctness between the selected samples and the classification results. The sequence of values that extend from the upper left corner of the matrix to the right corner is referred to as the diagonal which represent the number of pixels correctly classified for each categories.The following table(table 5.2) shows the error matrix generated from our classification results.

Table 5.2: Error matrix of the level 3 classification. The word comm in the table stands for commercial buildings, Grass L for grassland class, Art. S.F for artificial sport fields, Recr.a. for recreation areas, O. space for open space, Educat for educational buildings, Indust for industrial buildings, Swim.p for swimming pools, Parking p for parking plots, Resident for residential buildings and other non-res for other non residential buildings

Imag./Ref	comm.	Water	Trees	Roads	Grass L.	Art.S.F	Recr.a.	O. space	Educat.	Indust.	Swim. p.	Parking p.	Resident	Other $non_res.$
Commercial	16	0	0	1	0	0	0	0	0	2	0	0	0	2
Water	0	72	3	0	0	0	0	0	0	0	3	0	0	0
Trees	0	3	430	0	3	0	2	6	0	0	0	0	0	0
Roads	0	0	0	65	0	2	0	4	0	0	0	0	0	0
Grassland	0	1	3	0	60	0	12	3	0	0	0	3	0	0
Artificial S.F	0	0	0	2	0	18	0	2	0	0	0	3	0	0
Recreation areas	0	0	1	0	6	0	32	2	0	0	0	0	0	0
Open space	0	0	0	5	1	6	0	33	0	0	0	1	0	0
Educational	1	0	0	0	0	0	0	0	12	1	0	0	0	2
Industrial	1	0	0	0	0	0	0	0	2	12	0	0	0	1
Swimming pool	0	6	1	0	0	0	0	0	0	0	12	0	0	0
Parking plots	0	0	0	3	0	1	0	0	0	0	0	15	0	0
Residential	0	0	0	3	0	0	0	1	0	0	0	0	22	0
Other non residential	1	0	0	0	0	0	0	0	3	0	0	0	1	13

The error matrix above shows that roads and artificial sport fields classes is likely to be confused with open space in urban areas. Roads on the other hand tend to be confused with parking areas and parking areas are likely to be confused with open space.

The overall accuracy of our classification is computed as follows :

$$\text{Accuracy} = \left(\frac{the\ sum\ of\ diagonal\ values}{Total}\right) \times 100$$

$$\text{Accuracy} = \frac{812}{922} \times 100 = 88.069$$

The value found shows that the integration of additional information such as object shape, size, local and global context in the classification has improved the quality of the classification. To strength this observation and for a more objective assessment of our classification,the Kappa index was computed in order to evaluate the difference between a random classification results and our results. The Kappa index is a measure of the difference between the observed agreement between two maps and the agreement that might occur by chance matching of the two maps.

The first step in calculating the Kappa index is to define the expected accuracy which is derived from the product of row and column marginals. The row and column marginals have an important significance when estimating the chance agreement. They can be seen as areas occupied by each of the categories on the reference map and the image , and the chance of mapping any two categories is the product of their proportional extents on the two maps (Campbell,2006).

The following formula was used to calculate the Kappa index :

$$\text{Kappa index} = \frac{observed\ accuracy - chance\ agreement}{1 - chance\ agreement} (5.1)$$

(Ref : Lillesand and Kiefer,2000).

The computed expected agreement by chance is :

$$\text{Chance agreement} = \frac{Sum\ of\ diagonals}{sum\ of\ grand\ totals} (5.2)$$

Chance agreement $= \dfrac{219354}{815026} = 0.269$

The observed accuracy found was 0.88069 so the Kappa index is given as follows :

$$K= \frac{0.88069 - 0.269}{1 - 0.269} = 0.837 \quad (5.3)$$

The Kappa index value of 0.837 found revealed that the classification has achieved an accuracy of 83.7 percent better than would be an assignment of pixel to their respective categories by chance (Campbell,2006).

5.6 Discussion and conclusion

Some loss of details observed for some objects were attributed to object background such as shadow produced mostly by building and trees, causing an excess in object boundaries but not much could be done to prevent such limitations. Segmentation performed on large size objects was of good quality when the scale of 120 suggested by the investigation was used. The scale of 180 accurately extracted road network as a lower scale produced small road segments, which is not suitable for classification based on size and shape attributes. The segmentation of trees was satisfactory at this scale as the objects were organized into super-objects, easy to extract compare to single trees that could be confused with some individual buildings when classified based on size and shape attributes.

The high resolution aerial photograph used in this study offered spatial details that were exploited to separate non-residential buildings from residential buildings. However, the proximity of small adjacent parking lots in some residential areas altered building shapes. The extraction of water was challenging as most of small water bodies were merged with surrounding brighter objects such as individual parking lots in urban areas and even building roofs due to spectral similarities associated to their adjacency. The limitation was partly resolved with the use of fuzzy membership functions and taking into consideration size and shape information. Roads classification was satisfactory when using their low $\frac{Length}{Width}$ ratio, generally found smaller than 1.6. Grassland class was also accurately classified based on their high shape compactness which differentiated them from irregular tree patterns.

A few misclassification errors of open space as grass land were observed from the results.This confusion was certainly due to the spectral similarities between the two classes because of the presence of sand within some grassland areas. Parking plots were accurately identified especially the larger ones as the smaller individual parking lots were wrongly merged with open space class. Most of buildings were identified at the level 1 classification but predicted confusions revealed in chapter 4 with roads, parking lots and artificial sport fields were minimized by the classification sequence proposed in level 2. The entire tree cover was accurately classified and no major confusion occurred between this class and the others as they could perfectly be isolated in the spectral range between 10 and 95 in the red band. Conversely, some open space areas were confused with non residential buildings in the level 1 of classification probably due to spectral similarity between these classes.

The overall classification results found showed that multi-scale context based classification reduced the spectral variability within classes and ameliorated object recognition by extracting each object type at their suitable levels of aggregation. The analysis of classification results at level 1 based on pixel, size and shape information revealed a large number of salt and pepper outputs, but the integration of spatial, semantic, size, shape contexts as well as topology relations in higher levels, significantly reduced these limitations.

The different producer and user accuracies demonstrate the results above. Commercial buildings for instance had respective producer's and user's accuracy of 84.80 and 76.19 percent, roads had 82.27 and 91.54 percent, recreation parks 69.70 and 78.04 percent, educational buildings 70.08 and 75 percent, industrial buildings 80 and 75 percent, parking plots 68.18 and 78.94 percent.

Chapter 6

Conclusion and Future work

6.1 Conclusion

The aim of this dissertation was to characterize urban land use context for the rational of automated multi-scale classification of high resolution imagery. Urban land use information is relevant for scene understanding and analysis such as knowledge-based classification, that requires precise scene knowledge. This work tested the significance of incorporating urban scene contextual knowledge into a multi-scale classification system and the kappa index found at level 3 classification demonstrated the potential of performing classification based on scene descriptive knowledge, compared to a random assignment of pixels to their categories.

The statistics below (table 6.1) give a summary of the three level of classification results. As it was predicted, the use of full object context in level 3 classification improved the level 2 classification with 8.109 percent and the Kappa index increased from 74,84 to 83.70 between level 2 and level 3. For future study we intend to test our method on a different urban scene such as a rural area using high resolution images. This will assist in overcoming the challenge originating from the high spectral heterogeneity characterizing township scenes.

This study showed that combining context with multi-level classification can be an efficient approach of mapping complex urban scenes with a satisfactory accuracy. The success of the classification also relied on the optimal

Levels	Classification accuracy	Kappa index
1	63.67	56.07
2	79.96	74.84
3	88.069	83.70

Table 6.1: Classification statistics comparison between the three classification levels. Level 1 classification produced the lowest accuracy and Kappa index as it mostly relied on pixel information.

segmentation parameters selected using the technique we developed in chapter 5. Future related research will apply these methods of satellite data and preferably Ikonos or Geo-eye to exploit their spectral advantages.

Two main reasons motivated the proposal of such a classification system : (1) the proposition of a more objective method of optimal segmentation scale selection that minimizes segment internal heterogeneity and exploits inter objects heterogeneity. In the literature review, it was identified that previous methods used arbitrary segment areas instead of individual segment size measures,this created instabilities in segmentation results and decreased classifications' accuracies. The technique proposed here takes into account each size of segment generated by the segmentation to compute image variance. (2) the second reason for this work was to improve urban scene understanding by defining all the possible scene characteristics through a model in the form of high level feature space so that if any analyst desires to improve classification results of a residential land use for instance he can rely on the respective scene contextual information such as the general building size characterizing this type of areas, the shape of buildings, their spatial distances in respect to other objects within the scene, as well as their spatial organization if the objective is to distinguish a formal residential from informal residential scene for instance.

The merit of this research started with an investigation of the different techniques involved in identification and extraction of urban scene knowledge in order to build a model. Some of the mostly used segmentation and classification tools were also investigated and from this study, a segmentation approach based on multi-resolution algorithm was chosen in order to prepare a multi-scale context-based classification of the Cape Town urban scene. The advantage of a model-based classification is its robustness as it incorporates various objects real world attributes ranging from spectral characteristics to complex spatial organization of objects within the scene. The satisfactory final level 3 classification results showed that a discrimination of different urban land uses was possible based on the proposed contextual model. This demonstrates the effectiveness of the techniques used to extract the different high level features as well as the quality of these high level features. However, an analysis of the different classification error matrices and more particularly the level 3 classification error matrix showed that some misclassification occurred between certain urban classes. These limitations could originate from two hypotheses : (1) a limitation of the scene model which could not handle internal class variabilities. In fact, classes such as residential buildings, industrial, commercial, education buildings composing the built up category, exhibit similar spectral and shape properties and the fact that size of objects is highly affected by the segmentation process it is difficult to separate them based on size measures and the consequence can be some confusions between them. The other (2) hypothesis is the limitation in data involved in the segmentation process.

6.2 Future Work

Future research works are suggested in the following areas :

1. In order to minimize misclassification at objects boundaries, improvement is required not necessarily in methods of segmentation parameters selection but also in the use of ancillary data to supervise the segmentation. Data such as cadastral boundaries or digitized building foot prints are types of data that could segmentation results even when used with the best scale parameter. The use of high spectral resolution band such as the near-infra-red band might provide more object details than the visible light spectrum, and improve the segmentation of small objects such as some buildings, small parking plots

2. From the sixteen high level features investigated, only eight were tested. The use of other features such as building density could be of great support to separate residential buildings from non residential buildings. The use of textural features can also be investigated as it would have enabled the separation of building roofs from roads or parking plots.

3. High level features from other cities in South Africa or Africa could help building a more complete scene model that considers all the possible variabilities found in an urban scene. The more the samples the model

contains, the lower is the likelihood of finding misclassification due to certain similarities.

4. Even though a pre-classification simulation using techniques such as Bayesian networks in Genie smile can only predict the classification results, its use could have guided in giving more weight to certain high level features than others. This specific issue is currently being investigated for the purpose of the research article we are writing.

Bibliography

[1] Aber,J.S., Marzolff, I. and Ries, J.B., 2010. Small-format aerial photography: Principles, techniques and geosciences applications. Elsevier, Amsterdam, 268 p.

[2] Agouris P., Qyftakis S. and Setefanidis A., 2000. Uncertainty in image based change detection. Accuracy 2000, pp.1-8.

[3] Alejandra A.L.C., Mora F. and Escalante-Ramireaz B., 2004. Mapping and characterization of urban forest in Mexico. Proceedings SPIE 5239,522(2004).

[4] Aplin P., Atkinson, P. M. and Curttan, P. J., 1999a. Per field classification of land using the forthcoming very fine spatial resolution satellite sensors: problems and potential solutions. In P.M. Akinson and N.J. Tate (Eds),Advances in Remote Sensing and GIS Analysis,pp.219-239 (New York:John Wiley and Sons).

[5] Aplin P., Atkinson,P. M. and Curttan,P. J., 1999b. Fine spatial resolution simulated satellite sensor imagery for land cover mapping in the United Kingdom.Remote Sensing of Environment,68,pp.206-216.

[6] Alplin P. and Akinson P. M., 2001. Sub pixel land cover mapping for per field classification. International Journal of Remote Sensing,22,pp.2853-2858.

[7] Aria E., Saradjian M., Amini J., and Lucas C., 2004. Generalized co-occurrence matrix to classify IRS-1d image using Neural Network,In Proceedings of ISPRS Congress,Turkey,pp.117-123.

[8] Asmat A., Mansor S. and Hong W. T. ,2003. Rule-based classification for urban heat island mapping. 2nd Regional Conference, Marrakesh, Morocco, December 2-5, 2003.

[9] Awad M., Chehdi K. and Nasri A., 2007. Multi-component image segmentation using genetic algorithm and artificial neural network. Computer Journal of Geo-science and Remote Sensing Letters, Vol.4, No.4, pp.571-575.

[10] Awad M., 2010. An unsupervised artificial neural network method for satellite image segmentation. The International Arab Journal of Information Technology, Vol.7, No 2.

[11] Baatz M. and Shape A., 2000. Multi-resolution segmentation, an optimization approach for high quality multi-scale image segmentation. In: Strbl T.,Blaschke T., Griesebner G., eds. Angewandte Geographische informationsverarbeitung,XII. Beitrage zum AGIT Symposium Salzburg, Karlsruhe: Wichmann H. Verlag,pp.12-23.

[12] Bandyopadhyay S., 2005. Simulated Annealing Using a Reversible Jump Markov Chain Monte Carlo Algorithm for Fuzzy Clustering. IEEE Transactions on Knowledge and data Engineering. Vol.17,(4)(2005) 479-490.

[13] Barretti S., Del Bimbo A. and Vicario E., 2002. Spatial arrangement of colour in retrieval by visual similarity. Pattern Recognition 35 (2002) 1661-1674.

[14] Battle J., Casals A., Freixenet J. and Marti J., 2000. A review on strategies for recognizing natural objects in colour images of outdoors scene. Image and Vision Computing 18(2000) 515-530.

[15] Bauer T. and Steinnocher K., 2001. Per parcel land use classification in urban areas applying a rule-based technique. GeoBIT/GIS, 6, 24-27.

[16] Bellman R. and Zadeh E., 1987. Decision making in fuzzy environment, Fuzzy sets and applications, R. R. Yager, S. Ovchinnikov, R.M. Tong and H. T. Nguyen, Eds. New York: Wiley, pp.53-79

[17] Bellman R.E. and Zadeh L.A.,1987. Decision making in a fuzzy environment,Fuzzy Sets and Applications, R..R. Yager,S. Ovchinnikov,R.M.Tong,and H. T. Nguyen,Eds. New York: Wiley,1987,pp.53-79.

[18] Beneditksson J. A., Swain P. H. and Esroy O. K., 1990. Neural network approach versus statistical methods in classification of multi-source remote sensing data. IEEE transactions on Geoscience and Remote Sensing, Vol.25, No 4.

[19] Ben-Dor E., Levin N. and Saaroni H., 2001. A spectral-based recognition of the urban environment using visible and near infra-red region (0.4-1.1 micron meter): A case study over Tel-Aviv. International Journal of Remote Sensing 22, pp. 2193-2218.

[20] Benz U.C., Hofman P., Willhauck G., Lingenfelder I. and Heynen M., 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS ready information. Photogramm. Remote Ens., Vol.58, No.3/4, pp. 239-258.

[21] Bernsley M. J. and Barr S.L., 1997. Distinguishing urban land use categories in fine spatial resolution land cover data using a graph based,structured pattern recognition system. Computers, Environments and Urban Systems,21,pp.209-225.

[22] Beveridge J. R., Grffith J. Kohler R. R., Hanson A. R. and Riseman E. M.,1989. Segmenting images using localized histograms and region merging. International Journal of Computer Vision, 2(3), pp. 311-347.

[23] Binaghi E., 1992. Empirical learning for Fuzzy knowledge acquisition, In proc. 2nd International Conf. Fuzzy Logic Neural Networks, IIzuka, Japan, 1992, pp. 245-251.

[24] Binaghi E., Madella P., Montesano M. G. and Rampini A., 1997. Fuzzy contextual classification of multi-source remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, Vol.35, No.2.

[25] Binaghi E. Madella P., Montesano M.G., and Rampini A., 1997. Fuzzy contextual classification of multi-source remote sensing images. IEEE Transactions on Geoscience and Remote Sensing, 35, pp. 326-339.

[26] Binaghi E. and Rampini, A. 1993. Fuzzy decision making in the classification of multi source remote sensing data, Optic. Eng., Vol. 6 No.32,pp.1193-1203,1993.

[27] Binaghi E., Ventura A. D. , Grittini A. and Rampini A, 1993. Approximate reasoning techniques in remote sensing data analysis, In Proc. EGIS 93 Fourth European Conf. Exhibit. Geograph. Info. Syst. Genoa.Italy,1993,pp.503-511.

[28] Binaghi E. and Montesano M.G., 1994. A Fuzzy set based connectionist model with domain knowledge dependent learning rule, in Proc. IEEE IMCS Symp. Signal processing, Robotics Neural Networks, Lile, France, 1994, pp.472-475.

[29] Binaghi E., 1993. Fuzzy Logic-based tools for the acquisition and representation of knowledge in biomedical applications, in Fuzzy Reasoning Info., Decision Control System., MBISE series, A. N. Venetsanopulos and S. Tzafestas, Eds. Dordrecht,The Netherlands : Kluwer,pp.529-551.

[30] Bins L. S. A., Leila M. G. F., Guaraci J. E. and Fernando M. I., 1996. Satellite imagery segmentation : a region growing approach. Anais VIII Simposio Brasileiro de Sensoriamento Remoto, Salvador, Brasil, 14-19 Abril. INPE, p. 677-680.

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

[32] Bock M., Xofis P., Mitcheley J., Rossner G. and Wissen M., 2005. Object-oriented methods for habitat mapping at multi-scale. Case studies from Northern Germany and Wye Downs, UK. Journal of Nature Conservation 13(2-3),75-89.

[33] Bongiovani G. Cinque L., Leviialdi S. and Rosenfeld, A., 1993. Image segmentation by a multi resolution approach. Pattern Recognition, 26(12), pp. 1845-1854.

[34] Bonn F. and Escadafal R., 1996. La Teledetection appliquée aux sol. In: Bonn F. (Ed), Chapitre 3 Précis de Teledetection, Vol.l2 (Applications), PUQ/AUPELF.

[35] Bose B. and Grimson W. E. L, 2004. Improving object classification in far-field video. In proc. CBPR, 2004.

[36] Bouman C. A. and Shapiro M., 1994. A multi-scale random field model for Bayesian image segmentation. IEEE Transactions on Image Processing, 3(2), pp.162-177.

[37] Boyce R. R. and Clark W. A. V., 1964. The Concept of Shape in Geography. Geographical Review 54, 561-572.

[38] Brink A. B. and Eva H. D., 2008. Monitoring 25 years of land cover change dynamics in Africa. A sample based remote sensing approach. Applied Geography XXX: 1-12.

[39] Bruzzone L. and Carlin L., 2006. A Multi-level Context-Based System for Classification of Very High Spatial Resolution Images. IEEE Transactions on Geoscience and Remote Sensing, Vol. 44, No. 9.

[40] Buckner J., Pahl M., and Stahlhut O., 2000. GeoAIDA : A knowledge-based automatic image data analyser from remote sening data, CIMA 2001, Second International ICSC Symposium AIDA, June 19-22, Bangor, Wales, U.K., 2001.

[41] Busgeeth K., Van Den Bergh F., Whisken J. and Brits A. 2008. Potential Application of Remote Sensing in monitoring informal settlements in South Africa where complementary data does not exist. Proc. of SPIE Vol.74177147 OB-2.

[42] Chameau J. and Santamaria J. C., 1987. Membership functions I : Comparing methods of measurements. In Int. J. Approx. Reason, No. 1, pp. 287-317.

[43] Campbell J. B., 2007. Introduction to Remote Sensing. Fourth Edition. The Guilford Press, New York, USA.

[44] Chalifoux S., Cavayas F. and Gray J. T., 1998. Mapping guided approach for the automatic detection on Landsat Images of forest stands damaged by the spruce budworm. Photogrammetric Engineering and Remote Sensing, 64, pp. 629-635.

[45] Chang J. Y. and Cho C. W., 2002. Scene Analysis system using a combined fuzzy logic based technique. Journal of the Chinese Institute of Engineers, Vol.25, No. 3, pp. 297-307.

[46] Chang K-T., 2008. Introduction to Geographic Information Systems. Fourth Edition. New York, Mc Graw Hill.

[47] Carlotto M. J., 1997. Detection and analysis of change in remotely sensed imagery with application to wide area surveillance. IEEE Transactions on Image Processing, 6(1), pp. 189-201.

[48] Ceccato P. Connor S. J., Jeanne I. and Thomson M. C., 2005. Application of Geographical Information Systems and Remote Sensing technologies for assessing and monitoring malaria risk. Parassitologia 47 :81-96, 2005.

[49] Cheng H. and Bouman C. A., 1998. Trainable context model for multiscale segmentation. In Proceeding IEEE International Conference Image Processing 98, Chicago Il.

[50] Cheng H. and Bouman C. A., 2001. Multi-scale Bayesian segmentation using a trainable context model. IEEE Trans. Image Processing, Vol. 10, No.4.

[51] Chen H. Bouman C. A. and Allebah J. P., 1997. Multi-scale document segmentation, in Proceeding IS & T 50th Anniversary Conference, Cambridge, MA, pp. 417-425.

[52] Chen J., Delu P. and Mao Z., 2006. Optimum segmentation of simple objects in high resolution remote sensing imagery in coastal areas. Science in China series D Earth Science Vol.46 pp. 1195-1203.

[53] Chen X. L. , Zhao H. M., Li P. X. and Yin Z. Y. , 2006. Remote sensing image-based analysis of the relationship between urban heat island and land use/ land cover changes, Remote Sensing of Environment 104 (2006), pp. 133-146.

[54] Choi H. and Baraniuk R., 2000. Multi-scale document segmentation using wavelet domain hidden Markov models, in Proceeding IST/SPIE 12th Annual Symposium Electronic Imaging 2000, San Jose, CA.

[55] Choi H. and Richard G., 2001. Multi-scale Image Segmentation Using Wavelet Domain Hidden Markov Models. IEEE Transactions on Image Processing. Vol. 10, No 9.

[56] Chou Y. H., 1995. Spatial pattern and Spatial Autocorrelation. Spatial Information Theory Lecture Notes in Computer Science 988, 365-376.

[57] Cicala L., Poggi G. and Scarpa G., 2004. Supervised segmentation of remote sensing Multi-temporal Image Based on the Tree Structured Markov Random Field Model. 07803-87422/04/2004 IEEE.

[58] Cordley K. P. 1996. One meter satellite. Geo.Info. Systems. July 28, pp. 28-42.

[59] Cortijo F. J. and De La Blanca N. P. , 1998. Improving classical contextual classification. International Journal of Remote Sensing, No. 19, pp. 1591-1613.

[60] Cressie N., 1993. Statistics for Spatial Data.New York, Wiley.

[61] Dacey M. F., 1965. A Review of measures of contiguity for two and K Colour Maps. In spatial Analyses : A reader in statistical Geography, 479-495, edited by B. J. L. Berry and D. F., Marble Englewood Cliffs, NJ: Prentice Hall.

[62] Dare P. M., 2008. Small format digital sensors for aerial imaging applications. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. xxxvii, Part B1. Beijing.

[63] Dean A. M. and Smith G. M., 2003. An evaluation of per parcel land cover mapping using maximum likelihood class probabilities. International Journal of Remote Sensing, 24, pp. 2905-2920.

[64] De Jong S. M. and Van Der Meer F. D., 2005. Remote Sensing Image Analysis: Including the spatial domain. Netherland: Kluwer Academic.

[65] Donnay J. P., Barnsley M. J., Longley P. A., 2005. Remote Sensing and Urban Analysis. London, Taylor and Francis.

[66] Dey V., Zhang Y. and Zhong M.,2010. A Review of image segmentation techniques with remote sensing perspective. In : Wagner W., Szekely, B.(eds): ISPRS TC VII Symposium 100 years ISPRS, Vienna, Austria, 5-7. IAPRS, Vol. xxxviii, Part 7A.

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

[68] Duda R. O. and Hart P. E., 1998. Pattern classification and scene analysis, 1973.New York, John Wiley and Sons.

[69] Eiterer L. F., Facon J. and Menoti, 2004. Post envelop address block location by fractal-based approach. In 17th Brazilian Symposium on Computer Graphics and Image Processing, D. Coppersmith,Ed.,2004,pp.90-97.

[70] Elhadi E. M. and Zomrawi N., 2009. Change detection analysis by using Ikonos and Quickbird imageries. Nature and Science, 2009, 7(10).

[71] Espindola G. M., Camara G., Reis I. A., Bins L. S. and Montero A. M., 2006. Parameter Selection for Region-Growing Image Segmentation Algorithms using Spatial Autocorrelation. International Journal of Remote Sensing, in press, 2006.

[72] Everitt B. S., 1993. Cluster analysis.Third edition.1993. London: Halsted Press.

[73] Fan G. and Xia X. , 2001. A joint multi-context and multi-scale approach to Bayesian image segmentation. IEEE Transactions on Geoscience and Remote Sensing,39(12),pp.2680-2688.

[74] Flygare A. M., 1997. A comparison of contextual classification methods using Landsat TM. International Journal of Remote Sensing, 18, pp. 3835-3842.

[75] Foody G. M., 1997. Fully Fuzzy unsupervised Classification of Land Cover from Remotely Sensed Imagery with an Artificial Neural Network. Neural Computer and Applications (1997) 5: 238-247.

[76] Fosgate C. Krim H., Irvin W., Karl W. and Willsky A., 1997. Multi-scale segmentation and anomaly enhancement of SAR imagery. IEEE Trans. Image Processing, Vol. 6 pp.7-20

[77] Frauman, E., Wolff, E., 2005. Segmentation of very high spatial resolution satellite images in urban areas for segments-based classification. In: Proc. International Symposium Remote Sensing and Data Fusion Over Urban Areas and 5th Intern. Symposium Remote Sensing of Urban Areas, Tempe, USA, 14-16 March 2005.

[78] Gao Y., Kerle N., Mas J. F., Maathhuis B. H. P., Zhang X. M. and Van Dijk P. M., 2006. Comparison of pixel-based and object-oriented image classification approaches: a case study in a coal fire area, Wuda, Inner Mongolia, China. International Journal of Remote Sensing, 27, pp. 4039-4551.

[79] George L. W. P., Ben P. M. and Neal J. E., 2006. A comparison of methods for statistical analysis of spatial points patterns in plant ecology. Plant Ecol (2006) 187: 59-82.

[80] Goetz A. F.H., Vane G., Salomon J. E. and Rock B. 1985. Imaging spectrometry for Earth Remote Sensing. Science 228, pp. 1147-1153.

[81] Gond V., Freycon V. , Molino J. F., Brunaux O. , Ingrassia F., Joubert P., Pekel J. F., Prevost M. F., Thierron V., Trombe P. J. and Sabatier D., 2011. Broad scale spatial pattern of forest landscape types in the Guina Shield. International Journal of Applied Earth Observation and Geo-information, 13(2011) 357-367.

[82] Gong P. and Howarth P. J., 1992. Frequency based contextual classification and grey level vector reduction of land use identification. Photogrammetric Engineering and Remote Sensing, 58, pp. 423-437.

[83] Gonzalez R. C. and Wintz P., 1997. Digital Image processing. Second Edition. Boston, Addison Wesley Publishing Company.

[84] Goward S. N., Masek J. G., Williams D. L., Irons, J. R., Thompson R. J., 2001. The Landsat 7 Mission. Terrestrial research and applications for 21st century. Remote Sensing of Environment 78: 3-12.

[85] Gross M. H., Kock L. L. and Dreger A.,1994. Multi-scale image texture analysis in wavelet space. In Proceeding IEEE International Conference Image Processing 94, Austin,TX.

[86] Guindon B., Zhang Y. and Dillabaugh C., 2004. Landsat urban mapping on a combined spectral-spatial methodology. Remote Sensing of Environment 92, 218-232.

[87] Guo D., Atluri V. and Adam N., 2005. Texture based remote sensing segmentation. In : Proceedings of IEEE International Conference on Multimedia and Expo, pp. 1472-1475.

[88] Hamit Kok E., 2005. Developing an integrated system for semi-automated segmentation of remotely sensed imagery. Masters thesis. Middle East Technical University. The Graduate School of Natural and Applied Science.

[89] Haralick R. M. and Shapiro L. G., 1985, Image segmentation techniques. Computer vision, graphics, image processing, Vol.29, pp.100-132.

[90] Harintaka, Subaryono, Susanto A., Hartono, 2009. Assessment of low cost small format aerial photogrammetry for cadastral mapping. 7th FIG Regional Conference. Spatial Data Serving People : Land Governance and Environment Building the Capacity. Hanoi, Vietnam,19-22.

[91] Hepner G.F. and Chen C. M., 2002. Investigation of imaging spectroscopy for discriminating urban land covers and surface materials.In Proceedings AVIRIS Earth Science and Applications workshop, Palo Alto, CA, 2002.

[92] Hepner G. F., Houshmand B., Kulikov I., and Bryan N., 1998. Investigation of the integration of AVIRIS and IFSAR for urban analysis. Photogrammetric Engineering and Remote Sensing 6418, 813-820.

[93] Hepner, G. F. and Chen J., (2001). Investigation of imaging spectroscopy for discriminating urban land covers and surface materials. AVIRIS Earth Science and Applications Workshop, Palo Alto, CA (accessed: May , 2011) .

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

[95] Herold M. and Clarke K., 2002. A multi-scale framework for mapping and analysis of the spatial and temporal pattern of urban growth. Proceeding of 22nd EARSEL Symposium Geo-information for European wide integration, Prague.

[96] Herold M. M., Gardner B., H. and Roberts D., 2002. The spectral dimension in urban land cover mapping from high resolution optical remote sensing data. Proceedings of the 3rd Symposium on Remote Sensing of urban areas, Istanbul, Turkey.

[97] Herold M., Roberts D. A., Gardner M. E. and Dennisson P.E., 2004. Spectrometry for urban area remote sensing: Development and analysis of spectral library from 350 to 240 nm. Remote Sensing of Environment 9, 304-319.

[98] Herold M., Scepan J. , Muller A., Gunther S., 2002. Object-Oriented mapping and analysis of urban land use using Ikonos data. Proceedings of 22nd EARSEL Symposium Geo information for Europe wide integration, Prague, June, 2002.

[99] Herold M., Scepan J., Muller A., Gunther S., 2003. Object-Oriented mapping and analysis of urban land use/cover using Ikonos data. Geo-information for European wide Integration, Benes(ed) 2003, Millpress, Rotterdam.

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

[101] Hinz S. and Baumartner A., 2000. Road extraction in urban areas supported by context objects. International Archive of Photogrammetry and Remote Sensing. Vol.XXXIII, Part B3. Amsterdam.

[102] Huang F. and Qi Y., 2009. Object-Oriented land cover extraction in Changbai Natural reserve from Ikonos image. In proceedings of the International Conference on Geo informatics: GIScience in Change, University of Pekin, China,pp.1-4.

[103] Jain A. K. and Karu K., 1996. Learning texture discrimination masks. IEEE Transactions of Pattern Analysis and Machine Intelligence, Vol.18, No.2, pp. 195-205.

[104] Jain A. K. and Dubes R. C., 1988. Algorithm for clustering data. Prentice Hall, Englewood Cliffs, NJ.

[105] Jeansoulin R., Fontain Y. and Wermer F., 1981. Multi temporal segmentation by mean of fuzzy sets. In : Proceedings of seventh International Symposium on Machine processing of Remotely Sensed Data, LARS Purdue University, Layfette, pp. 336-339.

[106] Jeansoulin R., 1981. Multi temporal segmentation by means of Fuzzy Sets.Proceeding seventh International Symposium Machine Processing of Remotely Sensed Data with special emphasis on Range,Forest and Wetlands Assessment, Indiana, USA.

[107] Jianyu C., Delu P. and Zhihua M., 2006. Optimum segmentation of simple objects in high resolution remote sensing imagery in coastal areas. Science in China Series D: Earth Sciences 2006, Vol. 49 No. 11 pp.1195-1203.

[108] Jimenez, L. and Landgrebe, D. A. 1999. Hyper spectral data analysis and supervised feature reduction via projection pursuit, IEEE Transactions on Geosc. and Remote Sensing, 37, 6, pp. 2653-2667.

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

[110] Junxiang L. I., Liangjun D. A. Yujie W. and Yongchang S., 2006. Vegetation classification of East China with multi temporal NOAA-AVHRR data. Front. Bio. China(2006) 3:303-309.

[111] Karantzalos K. and Arglalas D., 2009. A Region based Level Set Segmentation for Automatic Detection of Man made Objects from Aerial and Satellite Images. Photogrammetric Engineering and Remote Sensing.

[112] Kardi T., 2007 . Remote sensing of urban areas: linear spectral unmixing of Landsat Thematic Mapper images acquired over Tartu (Estonia).Proceeding Estonia Acad. Sci. Biol. Ecol., 2007, 56, 1, 19-32.

[113] Kartikeyan B., Gopalkrishna B., Kalubarme M. H. and Majumder K.L.,1994.Contextual techniques for classification of high and low resolution remote sensing data. International Journal of Remote Sensing, 15, pp. 1037-1051.

[114] Kass M., Witkin A. and Terzopoulos D., 1987. Snakes : Active contour models. International Journal of Computational Vision, pp. 321-331.

[115] Keuchel J., Naumann S., Heiler M., and Siegmund A., 2003. Automatic land cover analysis for Tenerife by supervised classification using remotely sensed data . Remote Sensing of Environment ,86, pp. 530-541.

[116] Khedam R. and Belhadj-Aissa A., 2011. Classification of Multi-spectral Images using an artificial Ant-based algorithm. Cherifi H., Zain J.M. and El- Qawasmeh (Eds). DICTAP2011, PartI, CCIS166, pp.254-266.

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

[118] Kohonen T., 2001. Self organizing maps, Computer Journal of Springer series in Information Science, Vol.30, No.3, pp. 501-510.

[119] KoSir A. and Tasic J., 2002. Formal system based on fuzzy logic applied to digital image scene analysis. IEEE MELE XIN 2002, May 7-9,2002, Cairo, EGYPT.

[120] Krishnamachain S. and Chellapa R., 1997. Multi-resolution Gauss Markov random field models for texture segmentation,IEEE Trans.Image Processing,Vol.6,pp.251-267.

[121] Kuffer M., 2003. Monitoring the dynamic of informal settlements in Dar Es Salam by Remote Sensing: Explore the use of SPOT, ERS and Small Format Aerial Photograhy. CORP, 2003.

[122] Kunz D., Schilling K. J. and Vogtle T., 1998. A new approach for satellite image analysis by means of a semantic network. In Semantic Modelling and Extraction of Spatial Objects from Images and Maps (SMAT I97), edited by W. Forstner and L. Plumer, (Basel, Switzerland: Birkhauser Verlag), pp. 20-36.

[123] Kumar J.M., Garg P.K., and Khare D., 2008. Monitoring and modelling of urban sprawl using remote sensing and GIS techniques. International journal of applied earth observation and geo-information Vol.10, pp. 26-43.

[124] Lacave C. and Diez F., 2002. A review of explanation methods for Bayesian networks. The knowledge Engineering Review, Vol.17:2, 107-127, Cambridge University Press ,UK.

[125] Landgrebe D. A.,2003. Signal theory methods in multi-spectral remote sensing (Hoboken NJ: John Wiley and Sons).

[126] Landis J. and Koch G. G.,1977. The measurement of observer agreement for categorical data. Biometrics 33,159-174.

[127] Lee B.G., Tom and Carlotto M. J., 1986. A signal symbol approach to change detection. Proceed AAAI, August 1986.

[128] Lee D. R. and Salle G. T., 1970. A Method of measuring shape.Geographical Review, Vol.60, No.4 (October 1970), pp. 555-563.

[129] Lee J. and Wong D. W. S., 2001. Statistical Analysis with ArcView GIS. New York:Wiley.

[130] Lee J. Y., 2005. Automated Extraction of Road Networks from Ikonos Data in urban area. Proceedings of the 2005 IEEE International Geoscience Remote Sensing Symposium, Seoul, KOR, 2005, pp. 444-447.

[131] Lewinski S. and Zaremski K., 2004. Examples of Object-Oriented classification performed on High Resolution Satellite images. Miscellanea Geographica 11, 349-358.

[132] Lewinski S. and Bochenek Z., 2008. Rule based classification of SPOT imagery using object oriented approach for detailed land cover mapping.Proceedings of the 28th EARSel Symposium, Remote Sensing for a Changing Europe, 2-5 June, Istanbul, Turkey.

[133] Li H. T., Gu H. Y., Hans Y. S. and Yang J. H., 2008. An efficient multi scale segmentation for high resolution remote sensing imagery based on statistical region merging and minimum heterogeneity rule. International Workshop on Earth Observation and Remote Sensing Applications,11 pp.1-16.

[134] Li J., Gray R. M. and Olshen R. A., 2000. Multi-scale image classification by hierarchical modelling with two dimensional hidden Markov models. IEEE Trans. Inform. Theory, Vol.46, no.5 ,pp. 1826-1841.

[135] Lillesand T., Kieffer R. and Chipman J., 2008. Remote Sensing Image Interpretation. 6th Edition. New York: John Wiley & Sons.

[136] LIoyd C. D., Berberoglu S., Curran P. J. and Atkinson P.M.,2004. A comparison of texture measures for per field classification of Mediterranean land cover. International Journal of Remote Sensing,25, pp. 3943-3965.

[137] Lira J. and Maletti G., 2002. A supervised contextual classifier based on region growth algorithm. Computer and Geo-sciences 28(2002) 951-959.

[138] Liu Y., Guo Q. and Kelly M., 2008. A framework of region based spatial relations for non overlapping features and its application in object based image analysis. ISPRS Journal of Photogrammetry and Remote Sensing 63(2008) 461-475.

[139] Lu D., Mausel P., Batisstella M. and Moran E., 2004. Comparison of land cover classification methods in the Brazilian Amazon basin. Photogrammetric Engineering and Remote Sensing. 70(7) : 723-731.

[140] Lu D. and Weng Q., 2007. A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing, 28(5): 823-870.

[141] Lucieer A., Tsolmonsegerel O. and Stein A., 2004. Texture-based segmentation for identification of Geological units in Remotely Sensed Imagery. International Institute for Geo Information Science and Earth Observation (ITC), Netherlands.

[142] Lu D., Hetrick S. and Moran E., 2010. Land cover classification in a complex urban rural landscape with Quickbird imagery.Photogrammetric Engineering and Remote Sensing. Vol. 76 , No. 10, pp. 1159-1168.

[143] Lu D. and Weng Q., 2009. Extraction of urban impervious surface from Ikonos image. International Journal of Remote Sensing, 30(5) :1297-1311.

[144] Maceachren A. M. 1985. Compactness of Geographical Shape : Comparison and Evaluation of Measures. Geografiska Annaler. Series B. Human Geography, Vol. 67, No.1 (1985), pp. 53-67.

[145] Magneussen S., Boudewyn P. and Wulder M., 2004. Contextual classification of Landsat TM images to forest inventory cover types. International Journal of Remote Sensing, 25, pp. 2421-2440.

[146] Maktav D. C., Jürgens C., Siegmund A., Sunar F., Ebah H., Kalkan K., Uysal C., Mercan O.Y., Akar I., 2011. Multi-criteria Spatial Decision Support System for Valuation of Open Spaces for Urban Planning. 5th International Conference on Recent Advances in Space Technologies RAST 2011, Istanbul, Turkey, June 2011.

[147] Marangoz A. M., Oruc M., and salih K. B., 2004. Comparison of pixel-based and object-oriented classification using Ikonos imagery for automatic building extraction on Safranbolu Testfield. Proceedings of the XX th ISPRS Congress. Istanbul, 2004.

[148] Marinov M. and Zheliazkova I., 2005. An interactive tool based on priority semantic networks. Knowledge based systems 18 (2005) 71-77.

[149] Marques O., Berenholtz E. and Charvillat V., 2011. Context modeling in Computer Vision: Techniques, Implications and Applications. Multimed. Tolls. Appl (2011) 51: 303-339.

[150] Martha T. R., Kerle N., Van Westen C. J., Jetten V. and Kumar K. V., 2011. Segment Organization and Data-Driven Shareholding for Knowledge-based landslide Detection by Object-based Image Analysis. IEEE Transactions of Geoscience and Remote Sensing. Vol.49, No.12, December 2011.

[151] Massam B., H. and Goodchild, M. F., 1971: Temporal trends in the spatial organization of a service agency. Canadian Geographer, 15, 3: 193-206.

[152] Mather P. M., 2006. Computer processing of remotely Sensed Images: An Introduction . Third Edition. England, John Wiley & Sons.

[153] Maulik U. and Bandyay S., 2002. Performance Evaluation of some clustering Algorithms and Validity Indices. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol.24, (12) (2002) 1650-1654.

[154] Maulik U. and Saha I., 2009. Differential Evolution based Fuzzy Clustering Technique: Application to Satellite Image Segmentation. International Conference on Systemic, Cybernetics and Informatics 2009.

[155] Maxwell T. and Zhang Y., 2005. A fuzzy logic approach to optimization of segmentation of object oriented classification. In: Proceedings of SPIE 50th Annual Meeting and Photonics. San Diego,California, USA, 5909, pp.1-11.

[156] Mayaux P., Gond V. and Bartholome E., 2000. A near real time forest cover map of Madagascar derived from SPOT4 VEGETATION data. International Journal of Remote Sensing 21,3139-3144.

[157] Meinel G. and Neubert M., 2004. A comparison of segmentation programs for high resolution remote sensing data. International Archives of Photogrammetry and Remote Sensing, XXXV, 1097-1105.

[158] Mesev V., 2007. Integrating GIS and Remote Sensing. England : John Willey and Sons.

[159] Meng X., Wang L. and Currit N., 2008. Morphology-based building detection from Airborne Lidar data. Photogrammetric Engineering and Remote Sensing. Vol. 75, No 4, April 2009,pp. 437-442.

[160] Mc Keown D. M. , Harvey W. A. and McDermott J., 1985. Rule based interpretation of aerial imagery. IEEE Trans. Pattern Anal. Machine Intell. Vol.7, pp. 570-585.

[161] Michalis P., Dowman I. J., 2004. A regions model and DEM generation for SPOT HRS. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science 35(B1), pp .410-415.

[162] Misakova L., Jacquin A. and Gay M.,2006.. Mapping urban sprawl using VHR data and object oriented classification. EARsel Workshop of the SIG Urban Remote Sensing. Humbolt Universitat Zu Berlin,2-3 March 2006.

[163] Musaoghu N., Coskun M. and Kocabas V., 2005.Land use analysis of Beykoz Istanbul by means of satellite images and GIS. Available at nmusaoglu@ins.itu.edu.tr.(Accessed on June 2011).

[164] Moulin P. and Liu J., 1998. Analysis of multi resolution image denoising schemes using generalized Gaussian priors. In : Proceedings IEEE Signal Processing Int. Symposium,Time Frequency Time Scale Analysis, Pittsburg, PA, pp. 633-636.

[165] Narendra P. M. and Fukunaga K., 1977. A branch and bound algorithm for feature subset selection. IEEE Transactions in Computers. 26(9): 914-922.

[166] Navulur K., 2007. Multi-spectral Image Analysis using Object-Oriented Paradigm. Florida (USA), Taylor & Francis.

[167] Neibergall S., Loew A. and Mauser W., 2007. Object-Oriented Analysis of Very High Resolution Quickbird data for Mega City Research in Delhi, India. 2007 Urban Remote Sensing Joint Event.

[168] Neubert M. and Meinel G., 2003. Evaluation of segmentation programs for high resolution remote sensing applications.Proceedings of the joint ISPRS EARSel Workshop in High Resolution Mapping from Space 2003, Hanover, Germany.

[169] Nowak R., 1999. Multi-scale hidden Markov Markov model for Bayesian image analysis in Bayesian inference in Wavelet Based Models,p.Muller and B. Vidakovic, Eds. Berlin, Germany: Springer Verlag, pp. 243-266.

[170] Ojala T., Pietikainen M. and Harwood D., 1996 . A comparative study of texture measures with classification based on feature distribution. Pattern Recognition, Vol.29, No.1, 1996.

[171] Ojala T. and Pietikainen M., 1999. Unsupervised texture segmentation using feature distribution. Pattern Recognition,Vol.32, No.3, pp.477-486.

[172] Ojala T. Pietikainen M. and Maenpaa T., 2002. Multi-resolution grey scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(7) : 971-987.

[173] Pal N. R. and Bezdeck J. C., 1995. On cluster validity for the Fuzzy C-Means Model. IEEE Transactions on Fuzzy systems. Vol.3,(1995) 370-379.

[174] Pekkarinen A., 2002. A method for the segmentation of very high spatial resolution images of forested landscapes. International Journal of Remote Sensing 23 (14) pp.2817-2836.

[175] Peteri R. and Ranchin T., 2003. Urban street mapping using Quickbird and Ikonos images. in Proc. IEEE Int. Geosci. and Remote Sens. Symp., Toulouse, France, Jul. 2003, pp. 1721?1723.

[176] Pillay D. L. and Sebake M. D. , 2006. Mapping Post Apartheid Settlement growth patterns using Remote Sensing and GIS: A case of selected South African cities. African Association of Remote Sensing of the Environment (AARSE) 2006 Conference in Cairo, Egypt in October 2006.

[177] Pishva Davar, Kawai A. and Shiino T., 2000. Shape-based segmentation and colour distribution analysis with application to bread recognition. MVA 2000 IAPR Workshop on Machine Vision Applications, Nov. 28-30, 2000. The University of Tokyo,Japan.

[178] Poggi G., Scarpa G. and Zerubia J. B., 2005. Supervised segmentation of remote sensing images based on a tree structure MRF model. IEEE Transactions on Geoscience and Remote Sensing, 43(8), pp. 1901-1911.

[179] Porway J., Wang K., Yao B. and Zhu S. C., 2008. A Hierarchical and Contextual Model for Aerial Image Understanding. CVPR, pages 1?8, 2008.

[180] Pozzi, F., and Small C., 2002. Vegetation and population density in urban and suburban areas in the U.S.A., Proceedings of the. Third International Symposium of Remote Sensing of Urban Areas: pp. 489?496.

[181] Price K., 1996. Differential Evolution : A fast and simple numerical optimizer, NAFIPS 96. (1996) 524-527.

[182] Price K. and Reddy R., 1997. Change detection and analysis in Multispectral images. In Proceedings 5th International J. Conf. Artificial Intell., pp.619-625.

[183] Price K. and Storn R., 1996. Minimizing the real function of the ICEC 96 contest by Differential Evolution.IEEE International Conference on Evolutionary Computation(ICEC96 96).(1996) 842-844.

[184] Price K., Storn R., and Lampinen J., 2005. Differential Evolution: A Practical Approach to Global Optimization,Berlin, Springer.

[185] Pu R., Landry S. and Yu Q., 2009. Object based environment mapping with high resolution Ikonos imagery. ASPRS 2009 Annual Conference, Baltomore, Maryland, March 9-13,2009.

[186] Reed T. R. and Buf J. M. H. D., 1993. A review of recent texture segmentation and feature extraction techniques. Computer Vision Graphics and Image Processing : Image Understanding, 57(3), pp.359-372.

[187] Ridley H. M., Atkinson P. M., Aplin P., Muller J. P., Dowman I., 1997. Evaluating the potential of the forthcoming commercial U.S. High-resolution sensor imagery at Ordnance survey. Photogrammetric Engineering and Remote Sensing 63, 997-1005.

[188] Robert V., Maclntyre K., Keating J.,Trape J. F., Duchemin J. B.,Warren M. and Beir J. C.,2003. Malaria transmission in urban sub sahara Africa : American Journal of Tropical Medicine and Hygiene 68:169-176.

[189] Rocha A. F. and Yager R. R., 1992. Neural Networks and Fuzzy Logic. In Hybrid Architectures for Intelligent Systems,A. Kandel and G. Landholz, Eds. Boca Raton, Fl: CRC,pp.3-28

[190] Sadler G. J., Barnsley M. J. and Barr S. L., 1991. Information extraction from remotely sensed images for urban land analysis. Proceedings of the 2nd European Conference on geographical Information Systems. EGIS91, Brussels, Belgium, April (pp. 955-966) Utrecht: EGIS Foundation.

[191] Sadykhov R. and Mukashevisch M., 2008. Texture segmentation of satellite images by Neural Networks approach. The Second International Conference :Problems of Cybernetics and Informatics. September 10-12, 2008, Baku, Azerbaijan. Section 2, Intellectual Systems.

[192] Salman A. A. ,Ali A. E. and Mattar H. E., 2008. Mapping Land cover of Khartoum using Fuzzy Classification. Emirates Journal for Engineering Research, 13 (2), 27-43 .

[193] Salvador, S., and P. Chan., 2003. Determining the number of clusters segments in hierarchical clustering segmentation algorithms. Department of Computer Sciences Technical Report CS-2003-18, Florida Institute of Technology. Melbourne, Florida, USA.

[194] Shackelford A. K., Davis C. H., 2003. A combined Fuzzy pixel-based Object-Based approach for classification of High resolution Multispectral data over urban areas. IEEE transactions on geo-sciences and remote sensing vol. 4, No 10.

[195] Schumm S. N., 1963. Sinuosity of alluvial rivers on the Great Plains. Bulletin of the Geological Society of America,74 :1089-1100.

[196] Shen C., Li M., Chen J., Lu Y., 2010. Study on Urban Green Space Extraction from Quickbird imagery based on Decision Tree. In proceeding of the 2010 International Conference on Geoinformation, 18-20 June, 2010. Beijing, China.

[197] Shen, W., Jenerette, G. D., Wu, J., Gardner, R.H. (2004) Evaluating empirical scaling relations of pattern metrics with simulated landscapes. Ecography 27. 459-469.

[198] Singhal A., Luo J. and Zhu W., 2003. Probabilistic Spatial Context Model for scene Content Understanding. 1063-6919/03 (accessed on March 2011).

[199] Sirmacek B., d'Angelo P. and Reinartz P., 2010. Detecting complex building in panchromatic satellite images for Digital Elevation Model Enhancement. ISPRS Istanbul Workshop 1010 on Modelling of optical airborne and space borne sensors, WW1/4, Oct. 11-13, IAPRS Vol. XXXVIII.

[200] Small C., 2001. Estimation of urban vegetation abundance by spectral mixture analysis. International Journal of Remote Sensing, 22, 1305-1334.

[201] Small C., 2003. High spatial resolution spectral mixture analysis of urban reflectance. Remote Sensing of Environment, 88, 170-186.

[202] Smits P. C., Dellepiane S. G., Schowenge R. A., 1999. Quality assessment of image classification algorithm for land cover mapping : a review and a proposal for cost based approach. International Journal of Remote Sensing 20(8), 1461-1486.

[203] Song C., Woodcock C. E., Seto K. C., Lenney M.P. and Macomber S.A.,2001. Classification and change detection using Landsat TM data: When and how to correct atmospheric effects. Remote Sensing of Environment, 75, 230-244.

[204] Spirkovska L., 1993. A summary of segmentation techniques : NASA technical memorandum; 104022, California, USA.(assessed on June,2011).

[205] Storn R. and Price K, 1995. Differential Evolution.A simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces,Technical report TR 95-012, International Computer Science Institute,Berkley.

[206] Storn R. and Price K., 1997. Differential Evolution,A Simple and Efficient Heuristic Strategy for Global Optimization over continuous spaces, Journal of Global Optimization. Vol. 11,(1997) 341-359.

[207] Stow D. A. ,Lippitt C. D. and Weeks J. R ., 2010. Geographic object based delineation of neighbourhoods of Accra,Ghana using QuickBird satellite imagery. Photogrammetric Engineering and Remote Sensing.2010 August; 76(8): 907-914.

[208] Tabachnick B. G. and Fidell L. S. ,2001. Using Multivariate Statistics. Ally and Bacon, 4th edition, Needham Heights, MA. 2001.

[209] Tan Q., Liu Z. and Li X., 2009. Mapping urban surface imperviousness using SPOT Multispectral satellite images. IGARSS 2009.

[210] Taubenbosck, H., Pengler I., Schwaiger B., Cypra S., Hiete M., and Roth A., 2007. Multi-scale urban analysis of Hyderabad metropolitan area using remote sensing and GIS. Proceedings of 2007 Urban Remote Sensing Joint Event.

[211] Ter Braak C. J. F. and Prentice I. C., 1988. A theory of gradient analysis. Advances in Ecological Research. 18: 272-317.

[212] Therien C. W., Quatien T. F. and Dudgeon D. D.,1986. Statistical model based algorithm for image analysis.Proceedings IEEE,Vol. 74, No. 4.

[213] Thomas N., Hendrix C. and Congalton G. A., 2003. A comparison of urban mapping methods using high resolution digital imagery. Photogrammetric Engineering & Remote sensing, 2003, Vol.69, pp. 963-972.

[214] Ton J., Sticklen J. and Jain A. K., 1991. Knowledge based segmentation of Landsat images. IEEE transactions on geo-sciences and remote sensing, Vol.29, No.2.

[215] Unser M., 1995. Texture classification and segmentation using wavelet frames, IEEE Trans. Image Processing, Vol.4,pp.1549-1560.

[216] Venkatachalan V., Choi H. and Baraniuk R. G., 2000. Image segmentation using wavelet domain hidden Markov tree models. In Proceeding SPIE 14th International Symposium Aerospace/Defence Sensing Simulation, Control, Orlando, FL.

[217] Volpe F., 2003. Mapping information from space with Quickbird satellite data. Presented in Digital Earth 2003 Conference Brno (Czech Republic).

[218] Wang B. and Zhang L.,2003. Supervised texture segmentation using wavelet transform,in Proceedings of the 2003 International Conference on Neural Networks and Signal Processing, Vol.2, pp.1078-1082.

[219] Warton S. W., 1987. A spectral knowledge based approach for urban land cover discrimination.IEEE trans. Geosci. Remote sensing,Vol. GE-25, pp.272-282.

[220] Weng Q. and Quattrochi D. A., 2007. Urban Remote Sensing. New York: Taylor and Francis.

[221] Weszka, J. S., 1979. A survey of Threshold Selection Techniques. Computer Graphics and Image processing, Vol.l7, 1978, pp.259-265.

[222] William E. A. and Wentz E. A., 2008. Pattern Analysis Based on Type, Orientation, Size and Shape. Geographical Analysis 40(2008) 97-122.

[223] Wentz E. A, 2000. A shape definition of Geographic Applications Based on Edge,Elongation and Perforation. Geographical Analysis. Edinburgh, Scotland: Clustan Limited.

[224] Woodcock C. E. and Strahler A.H., 1987. The factor of scale in Remote Sensing. Remote sensing of Environment, 21(3), pp.311-332.

[225] Xiao X., Boles S., Liu J. , Zhuang D. and Liu M.,2002. Characterization of forest types in Northwest China, using multi temporal SPOT4 Vegetation sensor data. Remote Sensing of Environment, 82(2) : 335-348.

[226] Xu B., Gong P., Seto E. and Spear R., 2003. Comparison of grey level reduction and different texture spectrum encoding methods for land use classification using a panchromatic Ikonos image. Photogrammetric Engineering and Remote Sensing,69,pp.529-536.

[227] Yan G., Mas J. F., Maathuis B. H. P., Xiang Z., and Van Dijks P. M., 2006. Comparison of pixel based and object oriented image classification approaches: a case study in coal fire area,Wuda,Inner Mongolia, China. International Journal of Remote Sensing, Vol.27, No.18, pp.4039-4055.

[228] Yuan F., Sawaya K. E. ,Brian C., Loeffelholz C. and Bauer E., 2005. Land cover classification and change analysis of the Twin Cities (Minesota) Metropolitan Area by multi temporal Landsat remote sensing. Remote Sensing of Environment 98(2005) 317-328.

[229] Yin H. and Allison N., 1995. On the distribution and convergence of feature space in self organization maps.Computer Journal of Neural Computation, Vol.7, No.6 pp.1178-1187.

[230] Yue Y., Gong J. and Wang D., 2008. The extraction of water information based on SPOT5 image using object oriented method. Available from Yujuanlamei@126.com.[Accessed on June 2011].

[231] Zhang J., Rivard B. and Sanchez-Azofeifa A., 2005. Spectral unmixing of normalized reflectance data for the deconvolution of lichen and rock mixtures. Remote Sensing of Environment, 95, 57-66.

[232] Zhang L., Zhang J., Zhang D., Hou X. and Yang G.,2009. Urban Road Extraction from High resolution Remote Sensing Images Based on Semantic Model.PhD thesis , Capital Normal University,Beijing,China.

[233] Zhang P., Verma B. and Kumar K., 2003. Neural Vs Statistical classifier in conjunction with Genetic Algorithm Feature Selection in Digital Mammography. In Proceedings of IEEE Congress on Evolutionary computation CEC, Australia, pp.634-639.

[234] Zhang Y., 2006. A fuzzy logic approach to supervised segmentation for object oriented classification. ASPRS 2006. Annual Conference, Reno, Nevada,

[235] Zhou Z., Wei S.,Zhang X.,and Zhao X.,2007. Remote Sensing Image Segmentation Based on Self Organizing Map at Multiple Scale. In Proceedings of SPIE Geo informatics: Remotely Sensed Data and Information,USA,pp.122-126.